

## BACHELORTHESIS

# Analyse eines juristischen Entscheidungscorpus mit Methoden der Netzwerkforschung und Sprachtechnologie

vorgelegt von

Lennart Rönneburg

MIN-Fakultät

Fachbereich Informatik

Arbeitsbereich Language Technology (LT)

Studiengang: Software-System-Entwicklung

Matrikelnummer: 6920466

Abgabedatum: 18.01.2021

Erstgutachter: Prof. Dr. Chris Biemann

Zweitgutachter: Dirk Hartung

Betreuung: Dirk Hartung, Eugen Ruppert



# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
<b>2</b>	<b>Theorie</b>	<b>5</b>
2.1	Netzwerkforschung . . . . .	5
2.1.1	PageRank Algorithmus . . . . .	7
2.1.2	Chinese Whispers . . . . .	9
2.2	Sprachtechnologie . . . . .	10
2.2.1	LSTM . . . . .	12
2.2.2	Transformer Modelle . . . . .	13
<b>3</b>	<b>Related Work</b>	<b>15</b>
<b>4</b>	<b>Daten</b>	<b>19</b>
4.1	Korpus . . . . .	19
4.1.1	Netzwerkdefinition . . . . .	20
4.1.2	Erhebung der Daten . . . . .	21
4.1.3	Aufbereitung der Daten . . . . .	24
4.1.4	Software . . . . .	30
4.1.5	Statistiken des Korpus . . . . .	37
4.2	Trainingsdaten und Testdaten . . . . .	38
<b>5</b>	<b>Methodik</b>	<b>43</b>
5.1	Extraktion der Zitate . . . . .	43
5.2	Methoden der Netzwerkforschung . . . . .	46
5.2.1	Mikroebene . . . . .	47
5.2.2	Mesoebene . . . . .	49
5.3	Methoden der Sprachtechnologie . . . . .	50
5.3.1	Extraktion der Richter . . . . .	51
5.3.2	Klassifikation des Revisionsausgangs . . . . .	53
<b>6</b>	<b>Ergebnisse</b>	<b>57</b>
<b>7</b>	<b>Fazit</b>	<b>65</b>

## *Inhaltsverzeichnis*

**Hinweis:** Aus Gründen der Lesbarkeit wurde im Text auf die gleichzeitige Verwendung männlicher, weiblicher und diverser Sprachformen verzichtet. Sämtliche Personenbezeichnungen gelten für alle Geschlechter.

# 1 Einleitung

Ein Gesamtvolumen von über 1,2 Milliarden US-Dollar<sup>1</sup> und einzelne Investments von jeweils über 50 Millionen US-Dollar in Legal Technology Anbieter wie Kira Systems<sup>2</sup> oder Ironclad<sup>3</sup> in 2019 zeigen, welche immense wirtschaftliche Bedeutung technologiebasierte Rechtsdienstleistungen in den letzten Jahren gewonnen haben. Dahinter steht ein quantitatives Verständnis von Rechtsberatung, das zunehmend auf Methoden aus der Informatik setzt. Durch die steigenden technischen Möglichkeiten der Datenanalyse, auch von großen Datenmengen, ist dabei auf wissenschaftlicher Seite die quantitative Rechtswissenschaft stark in den Fokus geraten. Ziel ist es, mit Methoden der quantitativen Analyse von rechtswissenschaftlichen Daten Erkenntnisse über das Recht zu gewinnen und zu erkennen, wo das Rechtssystem verbessert werden kann.<sup>4</sup> Da Rechtswissenschaft bisher überwiegend normativ geprägt war, lassen sich vielfach neue Phänomene erstmals überhaupt nachweisen und bekannte Charakteristika des Rechtssystems messen. Ebenso entstehen neue Fragestellungen, deren Beantwortung vorher nicht möglich war.

Die quantitative Rechtswissenschaft steht jedoch auch vor einigen Herausforderungen. Viele Datensätze, welche sich für eine Analyse eignen, sind von geringer Größe, nur partiell verfügbar, unvollständig oder liegen in Formaten vor, die sich schlecht für weitere automatische Analysen verwenden lassen. Hier werden die klassischen Merkmale einer Analyse von *Big Data* (Volume, Velocity, Variety) häufig nicht vollständig erfüllt. Trotzdem lassen sich kleinere Datensätze bereits mit verschiedenen Methoden untersuchen und es lassen sich hier einige Erkenntnisse gewinnen und darstellen. Um eine empirische Analyse juristischer Daten, besonders mit größeren Datenmengen, bewältigen zu können, ist es jedoch notwendig, sich mit modernen computergestützten Werkzeugen und Methoden auseinanderzusetzen. Hier treffen die Disziplinen Informatik und Rechtswissenschaft aufeinander.

Die modernen, technischen Verfahren im Bereich der Informationstechnologie ermöglichen heutzutage noch nie dagewesene Optionen, um große Datenmengen zu untersuchen. Hier sind im Besonderen die Methoden und Techniken des maschinellen Lernens

<sup>1</sup><https://www.lawsitesblog.com/2019/09/at-1-1-billion-2019-is-a-record-year-for-legal-tech-investments-and-its-only-september.html>

<sup>2</sup><https://kirasystems.com/blog/one-year-reflections-bootstrapped-to-heavily-funded/>

<sup>3</sup><https://news.crunchbase.com/news/ironclad-raises-50m-to-help-companies-have-more-efficient-contracts/>

<sup>4</sup>vgl. [Bec20] S. 1

## 1 Einleitung

und der künstlichen Intelligenz zu nennen, welche maßgeblich für den Fortschritt in Datenanalysen mitverantwortlich sind. Eine Untersuchung von Gegenständen der Rechtswissenschaft mit künstlicher Intelligenz ist allerdings nicht neu. Bereits in den 1970er Jahren wurden unter anderem mit dem im den USA von L. Thorne McCarty entwickelten System *TAXMAN* erste Systeme umgesetzt, welche grundlegende Konzepte der künstlichen Intelligenz verwendeten. *TAXMAN* untersuchte damals Tatbestände aus dem Unternehmenssteuerrecht.<sup>5</sup> Dennoch zeigen exemplarisch Unternehmen wie *Leverson*<sup>6</sup> und *Luminance*<sup>7</sup>, welche in der letzten Dekade gegründet wurden und sich mit Machine Learning und künstlicher Intelligenz im Bereich der Rechtswissenschaft auseinandersetzen, dass das Thema zuletzt einen starken Aufschwung erfahren hat. Auch die Rechtswissenschaft profitiert hier von der Entwicklung.

Die Rechtswissenschaft, welcher die natürliche Sprache als eines ihrer wichtigsten Werkzeuge zugrunde liegt, ist vor allem auf Methoden der Sprachtechnologie angewiesen. Besonders die Entwicklung von modernen und effizienten Algorithmen, wie z.B. Transformer Modellen, welche in den letzten Jahren diese Forschung stark geprägt haben, ermöglichen mittlerweile relativ leicht umsetzbare Analysen natürlichsprachlicher Daten. Hiermit lassen sich nun außerdem ältere Analyseverfahren präzisieren und verbessern.

In dieser Arbeit werden die Entscheidungen der obersten deutschen Gerichte in der ordentlichen sowie der Verfassungs-, Verwaltungs-, Finanz, Arbeits-, Sozial und Patentgerichtsbarkeit mit Methoden aus der Netzwerkanalyse und Sprachtechnologie untersucht. Ziel ist es, unter Zuhilfenahme von modernen Datenanalyseverfahren verschiedene Instrumente aufzuzeigen, die eine Untersuchung von juristischen Daten auch mit größeren Datenmengen ermöglichen. Die Arbeit zeigt, welche Schritte für eine solche Untersuchung notwendig sind und führt schließlich einige Analysen durch, an denen die Durchführung nachvollzogen werden kann.

Im Rahmen dieser Analyse sollen die folgenden Fragen diskutiert werden: Was sind die zentralsten Entscheidungen der Gerichtsbarkeiten? Wie aussagekräftig ist ein Netzwerk von Entscheidungsdokumenten? Wo können mit modernen State of the Art-Methoden der Sprachtechnologie juristische Sachverhalte untersucht werden? Wie gut funktioniert eine Klassifikation des Revisionsausganges von Entscheidungen? Wie kann eine erweiterbare technische Infrastruktur geschaffen werden um die Daten für die Analyse vorzubereiten? Wie kann ein solcher Datensatz dauerhaft automatisiert aktuell gehalten werden?

Diese Arbeit beschäftigt sich nicht nur mit einer Beantwortung dieser Fragen, sondern erörtert auch verschiedene Metriken, welche für die Beantwortung solcher Fragen

---

<sup>5</sup>vgl. [Sus86] S. 175f. und [McC80] S. 23

<sup>6</sup>siehe <https://leverson.ai/artificial-intelligence-law/>

<sup>7</sup>siehe <https://www.luminance.com/>

realisieren können. Dabei verfolgt die Arbeit einen umfassenden Ansatz, sodass ein weiterer Schwerpunkt auf die Beschaffung und Vorbereitung der Daten sowie auf deren Speicherung, Analyse und Visualisierung gelegt wird. Diese Arbeit wird anhand einer Softwareinfrastruktur<sup>8</sup> demonstrieren, wie eine modulare technische Infrastruktur geschaffen werden kann, welche als Basis für Analysen und damit verbundene Applikationen dienen kann. Dieser Nutzen wird anhand einer der Software angegliederten Webanwendung deutlich, welche die Ergebnisse der Arbeit interaktiv darstellt. Die Arbeit adressiert zum einen Juristen, welche in dieser Arbeit einen Eindruck gewinnen, wie eine quantitative juristische Analyse von juristischen Daten angegangen werden kann. Zum anderen adressiert diese Arbeit gleichermaßen auch Informationstechnologen, welche in dieser Untersuchung eine Demonstration moderner Methoden der Netzwerkforschung und der Sprachtechnologie in dem spezifischen Anwendungsfall der Rechtswissenschaft erhalten.

Die vorliegende Arbeit ist in sechs Schwerpunkte unterteilt. Das Kapitel *Theorie* eröffnet die Arbeit mit einer grundlegenden Betrachtung der Themen Netzwerkforschung und Sprachtechnologie. Es werden hierzu bestehende Theorien und Konzepte vorgestellt, welche wichtig für die Durchführung dieser Arbeit sind, und deren Schlüsselbegriffe genauer erklärt. Anschließend beschäftigt sich das Kapitel *Related Work* mit themenverwandten Veröffentlichungen und setzt diese Arbeit in ihren forschungsrelevanten Kontext. Das Kapitel macht deutlich, wie sich diese Untersuchung von ähnlichen Untersuchungen unterscheidet und wo Parallelen zu erkennen sind. Das Kapitel *Daten* setzt die Untersuchung im Anschluss in die Praxis um. Hier wird der Prozess der Sammlung und Aufbereitung der Daten beschrieben und außerdem auf dabei auftretende Hindernisse eingegangen. Ebenso wird die Qualität der Daten kritisch betrachtet. Anhand einer Softwareanwendung zeigt dieser Teil auch, wie die gewonnenen Daten technisch gespeichert und für die Analyse vorbereitet werden. Im Rahmen dieser Arbeit wurde eine Webapplikation erstellt, welche in diesem Kapitel ebenfalls näher beschrieben wird. Anschließend wird im fünften Kapitel der Arbeit, *Methodik*, das methodische Vorgehen erläutert und an messbaren Ergebnissen präzisiert, warum die Ansätze dieser Arbeit für eine Untersuchung geeignet sind. Im Kapitel *Ergebnisse* werden die Daten schließlich analysiert und ausgewertet. Zuletzt werden im Kapitel *Fazit* die Ergebnisse der Arbeit zusammengefasst und das Potenzial für künftige Forschungen herausgestellt.

---

<sup>8</sup>[https://github.com/lroenneburg/bundesgerichte\\_law\\_corpus](https://github.com/lroenneburg/bundesgerichte_law_corpus)



## 2 Theorie

Die vorliegende Arbeit verwendet Methoden der Netzwerkforschung und Sprachtechnologie. Im folgenden Kapitel werden die zugehörigen technischen Grundlagen und Konzepte vorgestellt und genauer erläutert. Ferner werden die Konzepte in den Kontext ihrer übergeordneten Forschungsgebiete eingeordnet.

### 2.1 Netzwerkforschung

Die Netzwerkforschung bezeichnet die Untersuchung einer Menge von Objekten unter der zusätzlichen Betrachtung einer Menge von Beziehungen zwischen diesen Objekten.<sup>1</sup> Ein entscheidendes Fundament der Netzwerkforschung, vor allem um komplexe Netzwerke genauer analysieren zu können, bildet dabei die *Graphentheorie*. Die Graphentheorie ist ein Teilgebiet der Mathematik, welches die Beziehungen zwischen Objekten, unabhängig von ihrer Art oder ihrem Kontext, untersucht. Sie hat ihre Ursprünge im 18. Jahrhundert<sup>2</sup> und geht auf das sogenannte "Königsberger Brückenproblem" des Mathematikers Leonhard Euler zurück.<sup>3</sup> Durch die Möglichkeit, Probleme formal und abstrahiert beschreiben zu können, ergeben sich dabei zumeist ähnliche Lösungsansätze für Probleme mit eigentlich unterschiedlichen Bedeutungen.<sup>4</sup> Um die Komplexität eines Netzwerkes zu reduzieren und es formal zu spezifizieren eignet sich die Darstellung des Netzwerkes als *Graph*. Ein mathematischer Graph  $G = (V, E)$  besteht aus einer Menge von *Knoten*  $V$  (engl. *vertices*), welche die Objekte repräsentieren, und einer Menge von *Kanten*  $E$  (engl. *edges*), welche die Beziehungen zwischen den Objekten repräsentieren und somit die Knoten untereinander verbinden<sup>5</sup>. Jede Kante aus  $E$  gehört dabei eindeutig zu zwei Knoten, welche durch sie verbunden werden.<sup>6</sup> Die Kanten eines Graphen können *gerichtet* oder *ungerichtet* sein. Dabei ist die Menge  $E$  in ungerichteten Graphen definiert als Mengen der Form  $\{v_1, v_2\}$ , wobei  $v_1, v_2 \in V$ .

---

<sup>1</sup>vgl. [DJ10] S. 2

<sup>2</sup>In verschiedener Literatur wird das Jahr 1736 als das "Geburtsjahr" der Graphentheorie bezeichnet.

<sup>3</sup>vgl. [Bra10] S. 345

<sup>4</sup>vgl. [Bra10] S. 345

<sup>5</sup>Da das Gebiet der Graphentheorie in vielen unterschiedlichen Disziplinen Anwendung findet, sind im Laufe der Zeit auch unterschiedliche Begrifflichkeiten entstanden. So werden Knoten auch Ecken und im englischen auch *nodes* oder *points* genannt und Kanten sind im englischen auch unter den Begriffen *links* oder *lines* zu finden. In dieser Arbeit werden ausschließlich die Terme Knoten und Kanten verwendet.

<sup>6</sup>vgl. [IL14] S. 231

In gerichteten Graphen (auch *Digraphen* genannt) wiederum werden die Kanten in  $E$  als Paare statt Mengen definiert, da bei Paaren die Reihenfolge relevant ist, sodass die Kante eine eindeutige Richtung erhält. Ein solches Paar ist in der Form  $(v_1, v_2)$  definiert, wobei  $v_1, v_2 \in V$ . Ein Beispiel für einen ungerichteten Graphen ist: "Die drei Personen A, B und C kommunizieren miteinander". Hier verkörpern die drei Knoten jeweils eine Person, mit den zwischen ihnen liegenden Kanten, welche die Kommunikation abbilden. Ein Beispiel für einen gerichteten Graphen ist: "Person A sendet eine Nachricht an Person B, Person B sendet eine Nachricht an Person C und Person C sendet eine Nachricht an Person A". Hier bilden die drei Personen wieder jeweils einen Knoten, wobei die Kanten, welche die Kommunikation darstellen nun von einem Knoten auf einen jeweils anderen zeigen. Die Abbildung 2.1 illustriert diese Beispiele.

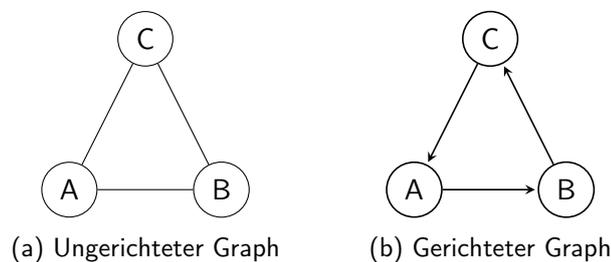


Abbildung 2.1: Illustrationen von Graphen am Beispiel

Zusätzlich gilt, wenn es in einem ungerichteten Graphen einen Knoten  $v \in V$  gibt, bei dem  $\{v, v\} \in E$  gilt oder in einem gerichteten Graphen  $(v, v) \in E$  gilt, so steht der Knoten mit sich selbst in Relation. Eine solche Beziehung nennt sich *Schleife* (engl. *loop*). Ein Graph ohne Schleifen wird *schleifenfrei* genannt<sup>7</sup>. In einer Beziehung von zwei Objekten können zudem unterschiedliche Beziehungsinstanzen auftreten. Ein Beispiel hierfür ist: "Person A sendet drei verschiedene Nachrichten an Person B". Da in der Graphentheorie nun alle Kanten als unterscheidbare Objekte definiert werden müssen, wird hier für die Kanten eine Multimenge verwendet. Solche Kanten nennen sich *Mehrfachkanten* (engl. *multiedges*). Besitzt ein Graph weder Schleifen noch Mehrfachkanten, wird dieser als *schlichter* oder *einfacher Graph* bezeichnet<sup>8</sup>. Bei der Darstellung von realen Sachverhalten als graphisches Modell können auch zusätzliche Bewertungen an Beziehungen geknüpft sein. Hierbei werden den Kanten Zahlen als Gewichtung zugeordnet, man spricht von *gewichteten Kanten* (manchmal auch *bewertete Kanten*) oder – bei einem Graphen mit ausschließlich gewichteten Kanten – von einem *gewichteten Graphen* (bzw. *bewerteten Graphen*)<sup>9</sup>. Ein Beispiel für gewichtete Kanten

<sup>7</sup>vgl. [IL14] S. 237 und [Bra10] S. 347

<sup>8</sup>vgl. [IL14] S. 237

<sup>9</sup>vgl. [IL14] S. 252

können dabei im Bezug auf die bereits genannten Beispiele Nachrichten sein, welche einer unterschiedlichen Priorität unterliegen. Abbildung 2.2 veranschaulicht diese Fälle.

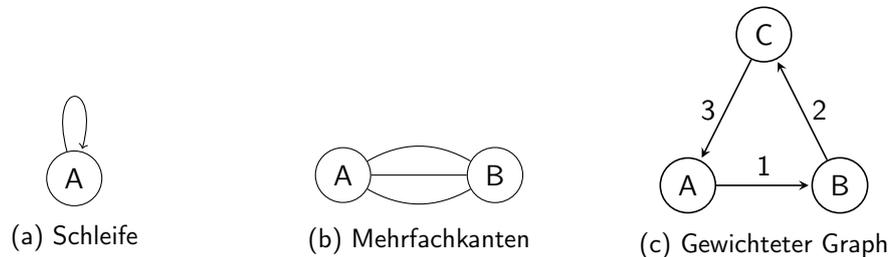


Abbildung 2.2: Illustration von Schleifen, Mehrfachkanten und Gewichtungen

Aufbauend auf den Grundlagen der Graphentheorie gibt es nun einige Methoden, welche bei der Untersuchung von Netzwerken hilfreich sein können.

### 2.1.1 PageRank Algorithmus

Innerhalb eines Netzwerkes kann es sehr hilfreich sein, zwischen der Relevanz der verschiedenen zu betrachtenden Objekte zu unterscheiden. So lässt sich zum Beispiel anhand der Anzahl der Beziehungen, welche auf ein Objekt verweisen, eine Rangordnung erstellen, in der Objekte, auf die häufig verwiesen wird, höher eingeordnet werden als Objekte, auf die nur selten oder gar nicht verwiesen wird. Diese Aufgabe erscheint für kleine Netzwerke noch recht überschaubar, kann jedoch mit steigender Größe eines Netzwerkes rasant sehr kompliziert werden. Eine Lösung, um auch in großen, verteilten Netzwerken die Objekte eines Netzwerkes nach ihrer Relevanz anordnen zu können, ist der *PageRank Algorithmus*. Ursprünglich in der Absicht, Webseiten im World Wide Web anhand von deren Verlinkungen zu gewichten, wurde der Algorithmus im Jahr 1999 von Lawrence "Larry" Edward Page und Sergey Brin, den Gründern von *Google Inc.*, entwickelt. Er bot die frühe Grundlage für ein Ranking in der Google-Suche. Der Algorithmus erhält dabei als Eingabe ein Netzwerk aus Objekten, welche gegebenenfalls miteinander in Beziehung stehen und erzeugt als Ausgabe eine Rangfolge der eingegebenen Objekte nach ihrer Gewichtung<sup>10</sup>. Um aus dem Netzwerk die Rangliste zu ermitteln, wird allen Objekten in dem gegebenen Netzwerk anhand der Struktur ihrer Beziehungen ein Gewicht zugeordnet. Dieses Gewicht wird als *PageRank* bezeichnet. Dabei gilt, je mehr Objekte eine Beziehung zu diesem Objekt haben, desto höher ist die Gewichtung dieses Objektes. Je nach der Höhe des PageRank der verweisenden Objekte wird diese Gewichtung zusätzlich beeinflusst.<sup>11</sup> Der PageRank Algorithmus

<sup>10</sup>vgl. [PBMW99] S. 2

<sup>11</sup>vgl. [MR11] S. 92

verfolgt dabei das *Random Surfer Model*, welches einen Nutzer simuliert, der zufällig innerhalb des Netzwerkes eine Beziehung zu einem anderen Objekt ansteuert.<sup>12</sup> Die Wahrscheinlichkeiten, mit denen der *random surfer* einen neuen Knoten ansteuert, ist dabei wiederum abhängig von dessen PageRank.

Mathematisch ist der PageRank wie folgend definiert: Sei  $u$  ein Objekt. Dann ist  $F_u$  die Menge aller Objekte, auf welche  $u$  selbst verweist und  $B_u$  eine Menge von Objekten, welche auf  $u$  verweisen. Zusätzlich sei  $N_u = |F_u|$  die Anzahl aller ausgehenden Verweise von  $u$ . Nun sei  $d$  ein sog. *damping factor* für die Normalisierung, wobei  $0 \leq d \leq 1$  damit die Gewichtung aller Objekte schlussendlich konstant bleibt.<sup>13</sup> Schließlich sei  $N$  die Menge aller Objekte im Graphen. Nun ist der PageRank  $R$  eines Objektes  $u$  folgend definiert:<sup>14</sup>

$$R(u) = \sum_{v \in B_u} \frac{R(v)}{N_v} * d + \frac{(1 - d)}{N} \quad (2.1)$$

Der PageRank  $R$  eines Objektes  $u$  berechnet sich anhand der auf  $u$  verweisenden Seiten  $B_u$ . Verlinkt ein Objekt  $v \in B_u$  nun auf  $N_v$  verschiedene andere Objekte, wird das Gewicht von  $v$  anteilig auf diese Objekte verteilt. Damit der *random Surfer* nicht auf einem Objekt landet, welches auf kein weiteres Objekt mehr verweist, wird dem Objekt zusätzlich ein Anteil des Gewichtes  $(1 - d)$  abgezogen und gleichmäßig auf  $N$ , also alle Objekte, umverteilt. Dies entspricht im Random Surfer Model einem Nutzer, welchem "langweilig" wird und welcher anschließend zu einem zufälligen Objekt im Netzwerk springt. Der Faktor  $d$  beträgt üblicherweise 0.85, kann jedoch auch innerhalb seines Wertebereiches angepasst werden.<sup>15</sup> Die Gleichung wird nun rekursiv so häufig ausgeführt, bis der PageRank der Knoten konvergiert.

Neben der Betrachtung der Relevanz von Objekten aus einem Netzwerk kann es außerdem sinnvoll sein, Objekte, die im Netzwerk ähnlich sind, zu gruppieren und somit Gemeinschaften zu bilden. Hier wird in der Informatik allgemein von *Clustering* gesprochen.<sup>16</sup> Im Bezug auf den jeweils vorliegenden Kontext muss dabei grundsätzlich der Begriff *ähnlich* genauer spezifiziert werden. In einem Netzwerk, wie es hier betrachtet wird, kann die Zusammengehörigkeit von Objekten über ihre Beziehungen ausgedrückt werden. Als Nachteil muss hier jedoch angeführt werden, dass für Objekte, welche keine Beziehung zueinander besitzen, auch kein Vergleich vorgenommen werden kann.<sup>17</sup>

Um zusammengehörige Objekte in einem Netzwerk sichtbar machen zu können, wird

<sup>12</sup>vgl. [PBMW99] S. 5

<sup>13</sup>vgl. [MR11] S. 92f.

<sup>14</sup>vgl. [MR11] S. 92

<sup>15</sup>vgl. [MR11] S. 93

<sup>16</sup>vgl. [Bie06] S. 73

<sup>17</sup>vgl. [Bie06] S. 73

hier also ein Verfahren benötigt, welches Cluster erkennt und separiert. Besonders für große Netzwerke wird dafür ein Clustering-Algorithmus benötigt, welcher sehr effizient arbeitet.

### 2.1.2 Chinese Whispers

Der Algorithmus *Chinese Whispers* ist ein Clustering-Algorithmus, welcher seinen Namen nach dem Kinderspiel "Stille Post"<sup>18</sup> erhalten hat. Der Kern des Spiels Stille Post besteht darin, sich innerhalb einer Gruppe von Teilnehmern gegenseitig eine Nachricht zuzuflüstern, welche sich durch das informelle Weitergeben der Nachricht durch die weiteren Teilnehmer stark verändern kann. Der Name Chinese Whispers lehnt an diesem Prinzip an. Das Verfahren versucht jedoch stattdessen Gruppen von Objekten in Netzwerken zu finden, welche für eine zentrale Nachricht stehen, die dann an ihre Nachbargruppen gesendet werden.<sup>19</sup> Chinese Whispers ist ein *harter* Clustering-Algorithmus und gehört den *partitionierenden* Clustering-Verfahren an. Er verfolgt eine *bottom-up* Vorgehensweise: Zu Beginn des Verfahrens wird jedem Objekt in dem Netzwerk eine eigene Klasse zugeordnet. Nun werden in einem Iterationsschritt die Objekte in einer zufälligen Reihenfolge einzeln betrachtet. Dabei erbt das betrachtete Objekt die Klasse von dem Nachbarobjekt, dessen Summe seiner Kantengewichte zum aktuellen Objekt maximal ist. Gibt es mehrere Objekte mit maximalem Kantengewicht, wird zufällig eines dieser Objekte ausgewählt. Wenn dies für alle Objekte im Netzwerk ausgeführt wurde, wird der Iterationsschritt wiederholt. Eine zusammenhängende Region stabilisiert sich nach und nach und wächst so lange, bis sie auf eine andere stabile Region trifft. Für den Fall, dass der Algorithmus konvergiert, sich innerhalb des Netzwerkes also die Klassen der Objekte nicht mehr verändern, ist die Berechnung abgeschlossen.<sup>20</sup> Dies muss jedoch nicht notwendigerweise passieren. Abbildung 2.3 veranschaulicht an einem Beispiel ein Netzwerk, welches nicht konvergiert. Die Farben Blau, Orange und

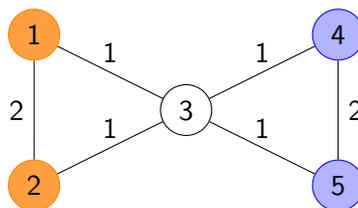


Abbildung 2.3: Nicht konvergierendes Netzwerk

Weiß stellen hier drei verschiedene Klassen dar. Objekt 3 kann hier nicht eindeutig einer Klasse zugeordnet werden, da die Kantengewichtung auf beiden Seiten gleich ist und

<sup>18</sup>engl. *Chinese Whispers* oder auch *Telephone game*

<sup>19</sup>vgl. [Bie06] S. 74

<sup>20</sup>vgl. [Bie06] S. 74

somit die Zuweisung zu einer der beiden Klassen zufällig erfolgt. Da sich dieser Prozess im nächsten Iterationsschritt wiederholt, entsteht hier eine Endlosschleife. Um solchen Endlosschleifen entgegen zu wirken, beendet Chinese Whispers seine Berechnung nach einer festgelegten Anzahl an Iterationsschritten automatisch, falls der Algorithmus vorher nicht bereits konvergiert. Die Anzahl der Iterationen, nach denen Chinese Whispers seine Berechnung beendet, hängt dabei vom *Durchmesser* des Netzwerkes ab.<sup>21</sup> Der Durchmesser ist definiert als der größte Abstand zweier Objekte im Netzwerk.<sup>22</sup>

Chinese Whispers bietet gegenüber anderen Clustering-Verfahren einige Vorteile: Im Gegensatz zu verschiedenen anderen Clustering Methoden setzt Chinese Whispers im Vorfeld keine Angabe über die Anzahl der zu bestimmenden Cluster voraus. Chinese Whispers ist außerdem mit einer Laufzeit  $O(n^2)$ , wobei  $n$  die Anzahl der Objekte im Netzwerk sei, sehr effizient. Der Algorithmus ist allerdings nicht-deterministisch. Führt man ihn also mehrfach auf ein identisches Netzwerk aus, so kann die Berechnung, durch die zufällige Auswahl von Knoten, in den Iterationen insgesamt zu unterschiedlichen Ergebnissen führen. Dies sollte als Nachteil berücksichtigt werden.

Wie eben demonstriert, können Verfahren wie der PageRank Algorithmus und Chinese Whispers eine systematische Betrachtung von Netzwerkstrukturen ermöglichen. Neben einer netzwerktechnologischen Betrachtung liegt dieser Arbeit noch ein weiterer Kern zugrunde: Die Sprachtechnologie.

## 2.2 Sprachtechnologie

Die Sprachtechnologie versteht sich grundsätzlich als die maschinelle Verarbeitung von menschlicher natürlicher Sprache, unabhängig davon, ob sie als gesprochene oder in Texten geschriebene Sprache vorliegt.<sup>23</sup> In der deutschen Sprache wird hier in einigen Zusammenhängen auch das Wort *Computerlinguistik* verwendet, wobei der Begriff Sprachtechnologie sich grundsätzlich näher mit dem Konstruieren von Sprachsoftware befasst.<sup>24</sup> Die Forschung der Sprachtechnologie besteht somit aus zwei wesentlichen Disziplinen, der Informatik und der Linguistik. Sprachtechnologischer Aufgaben können dabei sehr vielfältig sein. So zählen beispielsweise die Spracherkennung, das Übersetzen von Sprache oder Text, eine Rechtschreibprüfung oder das Erkennen von Entitäten innerhalb eines Textes zu diesem Feld. In der heutigen Zeit, in der Assistenzsysteme wie *Amazon Alexa*, *Apple's Siri* und der *Google Translator* allgegenwärtig sind, wird der Sprachtechnologie dabei mittlerweile eine immer höher werdende Bedeutung zugemessen. Alle diese Systeme verarbeiten Sprache, wenn auch in unterschiedlicher Form.

---

<sup>21</sup>vgl. [Bie06] S. 75

<sup>22</sup>vgl. [Bra10] S. 352

<sup>23</sup>vgl. [JM08]S. 1

<sup>24</sup>vgl. [CEE<sup>+</sup>09]S. 2

Obwohl sich die Sprachtechnologie sowohl mit gesprochener natürlicher Sprache als auch mit geschriebener natürlicher Sprache beschäftigt, wird der Fokus in dieser Arbeit auf geschriebene Sprache gelegt. Das Verarbeiten von natürlicher Sprache wird dabei als Teil des kognitiven Systems aufgefasst, in der sprachliches Wissen und nicht-sprachliches Wissen zusammen mit Denkprozessen und Handlungsentscheidungen in einem lernenden Prozess verbunden wird.<sup>25</sup> Es entstehen in der Sprachtechnologie also umfangreiche Überschneidungen mit dem *maschinellen Lernen*.

Maschinelles Lernen (engl. *Machine Learning*) bezeichnet im Kern den Prozess des Trainierens eines Computermodells um Wissen aus Daten, auf Basis von Erfahrung zu generieren.<sup>26</sup> Das Modell lernt aus Beispielen (*Trainingsdaten*) und versucht anhand dieser Beispiele eine allgemeine Erklärung abzuleiten. Mit dieser Erklärung soll dieses Wissen anschließend unbekannte Daten (*Testdaten*) einordnen können. Das Modell kann dabei voraussagend (engl. *predictive*) sein, um Vorhersagen über die Zukunft treffen zu können, oder deskriptiv (engl. *descriptive*) sein, um Wissen aus den Daten zu gewinnen. Ein Modell kann auch beide Ansätze verfolgen.<sup>27</sup> Machine Learning unterteilt sich dabei in zwei algorithmische Ansätze: Das überwachte Lernen (engl. *supervised learning*) und das unüberwachte Lernen (engl. *unsupervised learning*). Während beim überwachten Lernen das Modell mit Eingangsdaten trainiert wird, welche von einem Experten<sup>28</sup> annotiert wurden, gibt es beim unüberwachten Lernen keinen Experten, wodurch das Modell lediglich die Eingangsdaten als Ausgangspunkt für die Berechnung erhält und somit eigenständig Regularitäten in den Daten erkennen muss.<sup>29</sup> Die Forschung im Bereich des maschinellen Lernens ist dabei eng mit der Forschung von *künstlicher Intelligenz* (engl. *Artificial Intelligence*) verknüpft. Ein intelligentes System sollte schließlich auch in der Lage sein, sich an Veränderungen aus seiner Umgebung anzupassen.<sup>30</sup> Ein wichtiger Zweig der künstlichen Intelligenz, welcher auch in der Sprachtechnologie eine bedeutende Rolle spielt, sind dabei *künstliche neuronale Netze* (engl. *artificial neural network*). In künstlichen neuronalen Netzen, welche nach dem Vorbild von realen neuronalen Netzen im menschlichen Nervensystem entstanden sind, werden künstliche Neuronen erzeugt, um ein an das menschliche Gehirn angelehntes Netz zur Informationsverarbeitung zu erstellen. Dies dient nicht dem Zweck das menschliche Gehirn besser verstehen zu können, sondern verfolgt den Gedanken, dass künstliche neuronale Netze helfen können, effizientere Computersysteme zu konstruieren.<sup>31</sup>

---

<sup>25</sup>vgl. [CEE<sup>+</sup>09] S. 2

<sup>26</sup>vgl. [Nay18] S. 3

<sup>27</sup>vgl. [Alp20] S. 3

<sup>28</sup>in Literatur häufig *supervisor* oder *expert-annotator* genannt

<sup>29</sup>vgl. [Alp20] S. 11

<sup>30</sup>vgl. [Alp20] S. 15

<sup>31</sup>vgl. [Alp20] S. 267

Eine spezielle Form von neuronalen Netzen sind *rekurrente neuronale Netze* (engl. *recurrent neural networks*). Rekurrente neuronale Netze zeichnen sich dadurch aus, dass es innerhalb des Netzes verschiedene Arten von Rückkopplungen in den einzelnen Schichten des Netzes gibt. Ihnen stehen *Feed Forward Netze* gegenüber, welche keine Rückkopplungsschleifen besitzen.

Ein Verfahren, welches für künstliche neuronale Netze entwickelt wurde, ist *long short-term memory* (LSTM).

### 2.2.1 LSTM

Long short-term memory (deutsch: "langes Kurzzeitgedächtnis") ist ein Verfahren, welches die Leistungsfähigkeit von rekurrenten neuronalen Netzen in der künstlichen Intelligenz signifikant verbessert. Das Verfahren wurde im Jahr 1997 von Sepp Hochreiter und Jürgen Schmidhuber erstmalig vorgestellt. Rekurrente neuronale Netze, welche LSTM verwenden, sind in der Lage, sich effizient an frühere Erfahrungen zu erinnern. Es entsteht im Prinzip ein langanhaltendes Kurzzeitgedächtnis, welches dem Verfahren seinen Namen verleiht. LSTM greift dabei ein Problem von rekurrenten neuronalen Netzen auf: Je länger ein rekurrentes neuronales Netz trainiert wird, desto mehr Schichten entstehen und desto mehr Rückkopplungsschleifen besitzt ein Netz. In besonders tiefen Schichten des Netzes wird es dann allerdings besonders schwierig, Informationen effizient aufzufinden. Dadurch, dass das Netz die Informationen nicht mehr effizient aufrufen kann, werden sie gewissermaßen "vergessen". LSTM löst dieses Problem und schafft somit ein langanhaltendes Gedächtnis, welches effizient erreichbar ist.<sup>32</sup>

Ein LSTM Modul besteht aus verschiedenen Komponenten, sog. *Gates*. Das *input gate* steuert, welche Informationen als neue Werte in die Zelle fließen, und schützt die Zelle vor irrelevanten Eingaben. Das *output gate* wiederum schützt die anderen Zellen vor irrelevanten Eingaben der eigenen Zelle und steuert somit, welche in der Zelle vorhandenen Werte ausgegeben werden und welche nicht. Das *forget Gate* steuert, inwiefern Informationen in der Zelle verbleiben oder ob sie vergessen werden.

Der Aufbau von LSTM kann in der Abbildung 2.4 nachvollzogen werden.<sup>33</sup> Rekurrente Modelle wie LSTM oder GRU (*Gated Recurrent Units*), welches ähnlich zu LSTM funktioniert, wurden seit ihrer Entwicklung häufig für sprachtechnologische Aufgaben verwendet. Die Verfahren arbeiten dabei ihre Eingabe sequentiell ab. Später wurden die Methoden durch einen *Aufmerksamkeitsmechanismus* (engl. *attention module*) erweitert. Dieser Aufmerksamkeitsmechanismus ermöglicht es, die erhaltenen Eingabedaten im Kontext ihrer Umgebungsdaten zu betrachten. Dies können beispielsweise Wörter

---

<sup>32</sup>vgl. [HS97] S. 1

<sup>33</sup>Die Abbildung wurde unverändert entnommen aus [GMH13] S. 6646



## 2 Theorie

eine spezifische Aufgabe angepasst werden. Es wird hier von *fine tuning* gesprochen. Dabei erhalten sie als Eingabe einen Text in tokenisierter Form, wobei jedes Wort einen Token repräsentiert. Ihre Ausgabe geben sie ebenfalls in tokenisierter Form zurück. Dabei wird die Eingabesequenz von einer sog. *Embedding-Schicht* (engl. *embedding layer*) verarbeitet und in eine hochdimensionale Vektorrepräsentation überführt, um sie effizient weiterverarbeiten zu können. Beim Trainieren des Transformers wird diese Embedding-Schicht dann angepasst.

Transformer bestehen grundsätzlich aus einem Kodierer (engl. *encoder*) und einem Dekodierer (engl. *decoder*). Der Kodierer besteht dabei aus mehreren identischen Schichten, welche wiederum aus einem *self-attention* Modul und einem Feed-Forward Modul bestehen. Die Kodierer überführen dabei die eingegebene Sequenz in eine interne Repräsentation, welche die Semantik der Wörter abbildet. Auch die Dekodierer bestehen aus mehreren identischen Schichten, wobei hier in jeder Unterschicht ein zusätzliches Aufmerksamkeits Modul hinzukommt. Je mehr Kodiererschichten und Dekodiererschichten es gibt, desto länger kann auch die Eingabesequenz des Transformers sein<sup>35</sup>

Während rekurrente Netzwerke mit LSTM nur einen geringen Kontext, um die gerade bearbeitete Eingabesequenz betrachten konnte, ermöglichen Transformer Modelle eine weitaus größere Betrachtung des Kontextes.<sup>36</sup> Sie haben gegenüber LSTM außerdem den Vorteil, dass ihre Berechnungen wesentlich effizienter sind, einen geringeren Rechenaufwand benötigen und zudem gut skalierbar und parallelisierbar sind.

Die hier genannten Theorien werden im Rahmen dieser Arbeit eine Anwendung finden. Nun soll die Arbeit anhand der Betrachtung von themenverwandten Veröffentlichungen in ihren Forschungskontext eingeordnet werden.

---

<sup>35</sup>vgl. [VSP<sup>+</sup>17] S. 6000

<sup>36</sup>vgl. [VSP<sup>+</sup>17] S. 6000

### 3 Related Work

Obwohl die quantitative Untersuchung von rechtswissenschaftlichen Daten aufgrund der früher fehlenden technischen Möglichkeiten als eher jüngeres Forschungsgebiet eingestuft wird<sup>1</sup>, gibt es dennoch bereits verschiedene wissenschaftliche Beiträge, welche sich mit einer solchen Untersuchung auseinandergesetzt haben. Ein Blick in themenverwandte Veröffentlichungen macht deutlich, dass der Forschungsgegenstand der quantitativen bzw. empirischen Rechtswissenschaft ein wachsendes wissenschaftliches Interesse erfährt. Der folgende Abschnitt soll einen Eindruck davon vermitteln, wie in thematisch verwandter wissenschaftlicher Forschung die quantitative Analyse von Recht angegangen wird. Hierbei werden folgende Fragestellungen genauer betrachtet: Was passiert in quantitativen Untersuchungen von juristischen Daten im Bezug auf Netzwerke und Sprache? Wie und mit welchen Möglichkeiten können solche Untersuchungen betrieben werden? Dabei wird zunächst ein kurzer Überblick gegeben, was in der quantitativen Rechtswissenschaft passiert, und später genauer auf die Analyse durch netzwerk- und sprachtechnologische Mittel eingegangen. Wie kann eine Netzwerkanalyse in der Untersuchung von juristischen Daten behilflich sein und wo kann diese eingesetzt werden? Dazu werden folgende Fragen genauer erörtert: Welchen Nutzen hat Sprachtechnologie für die Untersuchung von Recht und in welchen Kontexten ergibt es Sinn eine Untersuchung durchzuführen? Abschließend beschäftigt sich das Kapitel mit der Visualisierung von quantitativen Forschungsergebnissen mit rechtswissenschaftlichen Bezug.

Was passiert bei der quantitativen Untersuchung von Recht? Typischerweise untergliedern sich quantitative rechtswissenschaftliche Untersuchungen in drei Schritte: Die Erhebung der Daten sowie deren Überprüfung und Aufbereitung, die Analyse dieser Daten und die Kommunikation der Auswertungen an die richtigen Adressatenkreise.<sup>2</sup> Um nun zu entscheiden, welche Daten für die Auswertung betrachtet werden sollen, ist allerdings zuerst die Entscheidung notwendig, welcher Gegenstand der Rechtswissenschaft nun genauer untersucht werden soll. Die rechtswissenschaftliche Staatstheorie unterscheidet drei Gewalten, die jeweils unter Berücksichtigung ihrer Besonderheiten quantitativ untersucht werden können<sup>3</sup>. Die Untersuchung der *Legislative* beschäftigt sich mit der Untersuchung von Gesetzestexten. Die Veröffentlichung [KCBH20] unter-

---

<sup>1</sup>vgl. [RHS<sup>+</sup>18] S. 2 und [Bec20] S. 1

<sup>2</sup>vgl. [Cou18] S. 380 und S. 386f. und [Har20] S. 17f.

<sup>3</sup>vgl. [Har20] S. 4

### 3 Related Work

sucht Gesetzestexte aus Deutschland und den USA und generiert daraus ein Netzwerk. Die Erzeugung des Netzwerkes erfolgt, ähnlich zu dieser Arbeit, über Zitate. Innerhalb des Netzwerkes der Veröffentlichung gibt es allerdings auch Kanten, welche die Strukturen der Normen abbilden, was in dieser Arbeit nicht weiter betrachtet wird. Im Gegensatz zu der Publikation, welche Referenzen von Gesetzestexten untersucht, beschäftigt sich diese Arbeit mit den Referenzen von Gerichtsentscheidungen verschiedener Gerichte. In der Veröffentlichung wird zudem ein internationaler Vergleich zwischen verschiedenen Rechtssystemen erhoben, wohingegen sich diese Arbeit auf die deutsche Rechtsordnung beschränkt. Neben der Legislative kann ebenso die *Judikative* betrachtet werden. Dies besteht dabei z.B. aus der Untersuchung von Gerichtsentscheidungen. Zuletzt kann auch die *Exekutive* untersucht werden, wobei hier beispielsweise eine Untersuchung von administrativen Daten, wie z.B. Verwaltungsverfahren möglich ist. Durch die Untersuchung von Gerichtsentscheidungen beschränkt sich der Rahmen dieser Publikation auf die Betrachtung von Daten aus dem Bereich der Judikative.

Im Rahmen von Datenuntersuchungen im rechtswissenschaftlichen Bereich gibt es bereits einige Veröffentlichungen, welche sich mit dem Schwerpunkt Netzwerkforschung beschäftigen. Veröffentlichungen wie [Cou19], [KCBH20] und [FJS<sup>+</sup>07] zeigen, dass Netzwerkforschung im Recht ein valider Weg ist, um neue Erkenntnisse zu gewinnen und erzeugen dabei eine grundlegende Vorstellung davon, welche Möglichkeiten für Netzwerkforschung im Recht bis heute bestehen und wie an eine solche Untersuchung systematisch herangegangen werden kann. Der Dissertation von Corinna Coupette [Cou19] liegt, ebenso wie dieser Arbeit, die Untersuchung eines Entscheidungsnetzwerkes zugrunde. Coupette überprüft jedoch auch, welche Objekte und Beziehungen sich für eine Untersuchung ebenso eignen können, und bewertet den Nutzen einer solchen Untersuchung. Im Unterschied zu dieser Veröffentlichung betrachtet Coupette allerdings lediglich Entscheidungen des Bundesverfassungsgerichtes (BVerfG). Coupette zieht hierbei einen weitaus größeren Datensatz an BVerfG-Entscheidungen heran. Aufgrund der Untersuchung von nur einer Gerichtsbarkeit werden im Netzwerk der Dissertation nur Zitate des BVerfG zu anderen Entscheidungen des BVerfG betrachtet, nicht jedoch Zitate zu anderen Gerichtsbarkeiten. Diesem Aspekt wird sich diese Arbeit annehmen. Coupette evaluiert in ihrer Untersuchung sowohl ein Entscheidungsnetzwerk mit gewichteten Kanten sowie mit ungewichteten Kanten. In dieser Arbeit wird lediglich ein Netzwerk mit ungewichteten Kanten betrachtet. Eine Perspektive über ein Zitationsnetzwerk zwischen verschiedenen europäischen Gerichten ermöglicht der Ansatz von M. Gelter und M. Siems [GS12]. Neben der Betrachtung von Netzwerken, welche aus juristischen Entscheidungen bestehen, so wie es in dieser Veröffentlichung vorgenommen wird, lassen sich allerdings noch weitere Gegenstände und Beziehungen im Recht untersuchen. Wie eine Netzwerkanalyse von sozialen Netzwerken aussieht, verdeutlicht [KS08]. Dabei untersucht der Artikel im genaueren die beruflichen- und

privaten Netzwerke von Sachbearbeitern der Richter des *Supreme Court of the United States* (Oberster Gerichtshof der Vereinigten Staaten) in einem Zeitraum von neun Jahren. Die Veröffentlichung arbeitet heraus, wie das soziale Netzwerk von Richtern dessen Gerichtsbarkeit beeinflusst. Des Weiteren lässt sich auch anstatt des Einflusses von sozialen Netzwerken, der Einfluss von Richtern durch ihre Nähe zu politischen Parteien quantitativ untersuchen, wie [EGS<sup>+</sup>17] zeigt. Hier wurden jedoch lediglich Richter des Bundesverfassungsgerichtes untersucht.

In verwandten Veröffentlichungen lässt sich ebenso betrachten, wie Sprachtechnologie in der Untersuchung von Recht behilflich sein kann. Das Untersuchen von Sprache im Bereich der Rechtswissenschaft ist schon früh in den Fokus der Forschung geraten. Der Artikel [Dal19] vermittelt ein grundlegendes Verständnis davon, wo sprachtechnologische Methoden in der Rechtswissenschaft eingesetzt werden können. Eine Erörterung verschiedener Methoden und Werkzeuge des *Natural Language Processing* (NLP) im Kontext der Jurisprudenz findet sich zudem in [Nay18]. Hier werden verschiedene Use-Cases beschrieben, in denen sich eine Forschung mit sprachtechnologischen Werkzeugen lohnt. Ein Beispiel für Untersuchungen, welche sich mit der Unterstützung der Entscheidungsfindung beschäftigen, ist [Fir12]. Hier wurde unter Zuhilfenahme von einem *Vector Space Model* ein Framework entwickelt, welches mit Methoden des begriffbasierten Text Minings den eben genannten Prozess unterstützt. Juristische Texte können auch in ihrer Semantik untersucht werden. Eine Arbeit, welche sich diesem Aspekt annimmt, ist die Dissertation von Bernhard Walzl [Wal18]. Die Arbeit widmet sich einer semantischen Analyse von Rechtstexten. Hier wird ein Framework entwickelt, in dem Dokumente mithilfe von Textanalyseverfahren um semantische Informationen ergänzt werden. Ebenso wie diese Arbeit verfolgt Walzl einen softwarearchitektonisch orientierten Ansatz, welcher sich auf die deutsche Rechtsordnung fokussiert.

Der Rahmen dieser wissenschaftlichen Arbeit befasst sich nur mit Entscheidungen des Bundesverfassungsgerichtes, der obersten Gerichtshöfe des Bundes, sowie des Bundespapentgerichtes. Einen weiteren Ansatz für die Untersuchung von Gerichtsentscheidungen liefert [ATPPL16]. Dieser unterscheidet sich allerdings von dieser Arbeit, da mit dem Europäischen Gerichtshofes für Menschenrechte (EGMR) eine andere Gerichtsbarkeit untersucht wird. In der Untersuchung wird zudem unter dem Einsatz eines *Support Vector Machine* Modells versucht, den Ausgang von Entscheidungen vorherzusagen.

Ein weiterer Ansatz, welcher sich der Vorhersage von Erfolg oder Misserfolg annimmt, hier allerdings von Berufungsentscheidungen im Steuerrecht, ist [WBS<sup>+</sup>17]. In dem Ansatz wird die Vorhersage mithilfe einer Klassifikation getroffen, was auch in [RHS<sup>+</sup>18] eine Rolle spielt, hier allerdings angewendet wird, um den Revisionserfolg/ -misserfolg von Entscheidungen des Bundesgerichtshofes (BGH) zu bestimmen. Im Zuge der Extraktion von juristischen Entscheidungsträgern aus den Entscheidungsdokumenten in

### 3 Related Work

dieser Arbeit, zeigt [TYS<sup>+</sup>20] zudem, wie auf Basis einer Deep-Learning basierenden Methode Richter als eine Entität extrahiert werden können. Auch die Publikation [YB19] diskutiert eine Methodik, welche Entitäten unter Zuhilfenahme von Deep Learning Algorithmen extrahiert. Der Ansatz in [CN16] demonstriert ein Named Entity Recognition Modell anhand eines bidirektionalen LSTM in einem rekurrenten neuronalen Netzwerk. In einer Vorarbeit [RHS<sup>+</sup>18] wird außerdem eine Heuristik verwendet, um Revisionsentscheidungen zu klassifizieren. Diese Arbeit soll mit modernen Transformer-Methoden (ELECTRA, [CLLM20]) die Klassifikationsergebnisse verbessern.

Um nach der Analyse den richtigen Adressaten zu ermöglichen die Ergebnisse nachzuziehen und zu untersuchen, ist eine entsprechende Visualisierung notwendig. Dies ist zuletzt auch Gegenstand dieser Arbeit. Daher wird im Verlauf der Arbeit auch ein Ansatz zur Visualisierung der Forschungsergebnisse gezeigt. Zudem realisiert [LFM<sup>+</sup>18] am Projekt *EUCaseNet* eine Lösung, um eine Netzwerkanalyse im Kontext von Recht graphisch darzustellen. In *EUCaseNet* wird ebenso wie in dieser Arbeit ein Zitationsnetzwerk aufgebaut, jedoch, wie dem Namen bereits zu entnehmen ist, auf Basis europäischer Fälle. Allerdings werden im Projekt auch Referenzen auf Literaturbeiträge mit in die Analyse eingeschlossen. Dies ist in dieser Untersuchung nicht der Fall.

Der Blick auf verwandte Arbeiten macht deutlich, wie vielfältig die Rechtswissenschaft mit Methoden der Informatik untersucht werden kann. Die folgenden Kapitel sollen nun die Aspekte Netzwerkforschung und Sprachtechnologie im Recht aufgreifen und verschiedene Untersuchungen an Daten der rechtswissenschaftlichen Praxis durchführen.

## 4 Daten

Quantitative Untersuchungen verstehen Daten als ihre Grundlage. Das folgende Kapitel klärt, welche Daten für die Untersuchungen dieser Arbeit benötigt werden, beschreibt den Prozess der Gewinnung dieser Daten und stellt anschließend das daraus resultierende Korpus vor. Zuletzt wird außerdem ein Ausschnitt der Daten beschrieben, welcher als Datensatz zum Trainieren und Testen von maschinellen Lernmodellen verwendet wird.

### 4.1 Korpus

Um innerhalb eines Netzwerkes Entscheidungsdaten analysieren zu können, müssen zuerst die für die Untersuchung nötigen Daten gesammelt und aufbereitet werden. Dabei sind für die Sammlung der Daten bereits die folgenden Entscheidungen notwendig: Welche Daten sollen innerhalb des Netzwerkes für das Experiment betrachtet werden? Aus welchen Quellen können diese Daten bezogen werden? Abschnitt 4.1.1 erklärt, wie das Netzwerk dieser Arbeit definiert wird und begründet diese Entscheidung. Im Abschnitt 4.1.2 wird dann genauer beschrieben, welche Daten aus welchen Datenquellen für das Experiment in Frage kommen und warum sich für diese Quellen entschieden wurde. Der Abschnitt 4.1.3 thematisiert, welche Hindernisse bei der Aufbereitung der Daten auftreten und diskutiert verschiedene Lösungswege. Damit auf die gesammelten Daten anschließend Analysen ausgeführt werden können, soll eine geeignete Infrastruktur bereitgestellt werden, welche die Daten persistent speichert, schnell abrufbar macht und eine flexible Weiterverarbeitung ermöglicht. Abschnitt 4.1.4 erläutert dieses Thema genauer und demonstriert dabei außerdem eine Möglichkeit, um den Datenbestand fortlaufend aktuell zu halten. Um einen genaueren Einblick über den Aufbau des Systems zu vermitteln, wird dieser Abschnitt außerdem die Architektur und den Datenfluss einer Entscheidung innerhalb des Systems vorstellen. Weiterhin wird in Abschnitt 4.1.5 das Korpus anhand verschiedener Parameter vorgestellt. Der Abschnitt 4.2 schließt das Kapitel mit der Betrachtung eines generierten Datenausschnittes für das Trainieren und Testen von maschinellen Lernalgorithmen.

### 4.1.1 Netzwerkdefinition

Um innerhalb des rechtswissenschaftlichen Korpus eine netzwerkanalytische Untersuchung durchzuführen, muss zu Beginn entschieden werden, welche Akteure im Netzwerk als Knoten und welche Beziehungen als Kanten zwischen den Knoten des dahinterstehenden Graphen fungieren sollen. Auf dieser Basis können dann die dafür nötigen Datenquellen evaluiert werden. Es erscheint sinnvoll von einem Dokumentennetzwerk auszugehen, in dem jedes Entscheidungsdokument als Knoten und die Zitate, welche innerhalb eines Entscheidungstextes auf eine andere Entscheidung verweisen, als Kanten zu modellieren. Da hier ein asymmetrischer Verweis von nur einer Entscheidung auf eine andere besteht, liegt es nahe, das Netzwerk durch einen gerichteten Graphen abzubilden. Es kann nun vorkommen, dass eine Entscheidung mehrfach innerhalb des Entscheidungstextes auf eine andere Entscheidung verweist. An dieser Stelle ist die Entscheidung notwendig, ob das Netzwerk solche Mehrfachzitate abbilden soll oder nicht. Der Fokus dieser Arbeit liegt primär darauf, Entscheidungen zu untersuchen, welche von vielen unterschiedlichen anderen Entscheidungen zitiert werden. Innerhalb dieses Netzwerkes, werden somit nur einfache Referenzen einer Entscheidung zu einer anderen Entscheidung betrachtet. Es ist also nicht von Bedeutung, wie oft eine Entscheidung  $A$  eine Entscheidung  $B$  zitiert, sondern lediglich, dass Entscheidung  $B$  von Entscheidung  $A$  zitiert wird. Die Zitate werden binär gezählt und statt Mehrfachkanten zu verwenden, werden sie als eine einzelne Kante dargestellt. Da eine Gerichtsentscheidung sich außerdem nicht selbst zitiert, kann in der Netzwerkdefinition von einem schlichten Graphen gesprochen werden. Ein solches Netzwerk soll ermöglichen, zusammenhängende Entscheidungen schnell erkennen zu können und dabei im besten Fall verschiedene Entscheidungscluster bilden zu können, welche Urteile bündeln, die sich untereinander direkt oder indirekt referenzieren. Ein weiterer Aspekt einer solchen Netzwerkanalyse wird sein, die innerhalb des Korpus am häufigsten zitierten Entscheidungen zu bestimmen. Es wird jedoch im Hinblick auf die Datengrundlage des Korpus ein weiterer Gesichtspunkt deutlich, welcher zunächst beantwortet werden muss. So werden in Entscheidungen der obersten Gerichtshöfe nicht nur andere Entscheidungen dieser, sondern auch Entscheidungen von Gerichten außerhalb des Korpus, z.B. Landes- oder Oberlandesgerichte zitiert. Es besteht also die Möglichkeit, ein *geschlossenes Netzwerk* zu untersuchen, somit also nur Zitate zu Entscheidungen zu betrachten, welche auch Teil des Korpus sind. Ebenso kann ein *offenes Netzwerk* betrachtet werden, welches auch Zitate zu Entscheidungen enthält, die nicht im Korpus enthalten sind. Ein geschlossenes Netzwerk bietet dabei den Vorteil, dass die Menge der Entscheidungen im Netzwerk nicht rekursiv durch die Entscheidungen außerhalb des Netzwerkes erweitert werden muss.<sup>1</sup> Somit haben alle Entscheidungen im Netzwerk ihren Ursprung

---

<sup>1</sup>Rekursiv meint hier die Erweiterung des Netzwerkes durch die direkt zitierten Entscheidungen, sowie dessen zitierte Entscheidungen, usw..

im Korpus. Jedoch werden in einem geschlossenen Netzwerk auch wichtige Entscheidungen außerhalb des Korpus nicht betrachtet, obwohl sie möglicherweise sehr häufig zitiert werden. Bei einem offenen Netzwerk ist nun zu entscheiden, wie mit Referenzen zu Entscheidungen außerhalb des Korpus umgegangen wird. In dieser Untersuchung wird die Menge der Zitate in einem offenen Netzwerk betrachtet, um alle im Netzwerk enthaltenen Kanten einschließen zu können und keine wichtigen Informationen zu verlieren. Es wird im Korpus allerdings davon abgesehen die Datenbasis durch zitierte Entscheidungen zu erweitern, da dies die Verwendung mehrerer verschiedener Datenquellen bedeuten würde. Innerhalb des offenen Netzwerkes werden lediglich Entscheidungen außerhalb der Datenbasis im Netzwerk betrachtet, welche direkt von einer Entscheidung aus dem Korpus zitiert wurden.

### 4.1.2 Erhebung der Daten

Um juristische Daten zu untersuchen, bieten sich grundsätzlich zwei Möglichkeiten: Das Untersuchen von gesamten juristischen Fallakten und das Untersuchen von den in ihnen enthaltenen Gerichtsentscheidungen<sup>2</sup>. Die Analyse von juristischen Fallakten stellt den sozialen Kontext des Verfahrens in den Mittelpunkt und hat damit eher einen sozial- als rechtswissenschaftlichen Fokus. Der Zugriff auf die Daten von Fallakten, vor allem in größeren Mengen, ist allerdings schwierig umsetzbar. Eine Analyse von Gerichtsurteilen eignet sich somit vor allem in einer quantitativen Hinsicht eher, da diese auf verschiedenen Plattformen auch in größeren Mengen veröffentlicht werden. Um für eine Untersuchung juristische Entscheidungsdaten zu beziehen, bieten sich aktuell folgende Quellen an:

#### Juris

Das *juristische Informationssystem (Juris)*<sup>3</sup> ist ein von der *juris GmbH* zur Verfügung gestelltes Online-Portal, welches den Zugang zu einer Datenbank für Rechtsinformationen ermöglicht. Ursprünglich in den 1980er Jahren von der Bundesregierung initiiert, war Juris zuerst dem Bundesministerium der Justiz angehörig, wurde später jedoch als Gesellschaft mit beschränkter Haftung (GmbH) ausgegliedert, welche heute in einer Mehrheit der Bundesrepublik Deutschland und in unterschiedlich großen Teilen dem Bundesland Saarland, dem deutschen Anwaltverein, der Bundesrechtsanwaltskammer und weiteren Verlagen gehört. Um Zugang zu der Datenbank zu erhalten, bietet Juris ein abonnementbasiertes Geschäftsmodell an.<sup>4</sup> Die Datenbank von Juris umfasst neben Gesetzen, Vorschriften, Zeitschriften und weiteren Dokumenten mehr als 1,5 Millionen

---

<sup>2</sup>vgl. [Har20] S. 5

<sup>3</sup>siehe <https://www.juris.de/>

<sup>4</sup>vgl. [Har20] S. 8

Gerichtsentscheidungen.

Eine Auswertung der Entscheidungsdaten von Juris stehen dabei allerdings zwei Probleme im Weg: Neben der Möglichkeit eines kostenpflichtigen Abonnements stellt Juris nur geringfügige Möglichkeiten zur Verfügung, um ihre Datenbank wissenschaftlich quantitativ untersuchen zu können. So hat es in der Vergangenheit bereits Fälle gegeben, in denen Wissenschaftler Zugang zu Teilen der Datenbank erhielten. Jedoch gibt es hierfür kein standardisiertes Verfahren, wodurch Anfragen einzeln bearbeitet und der Zugang erst mit Verzögerungen eingerichtet werden kann. Juris erwartet (aus wirtschaftlich nachvollziehbaren Gründen) zudem für den Zugang zu größeren Datenmengen, dass eine Verschwiegenheitsvereinbarung unterzeichnet wird, welche zusätzlichen Aufwand bei der Veröffentlichung der Ergebnisse der Untersuchung bedeutet.<sup>5</sup>

Die Gerichte sind außerdem frei in der Entscheidung, welche Urteile sie Juris zur Aufnahme in die Datenbank übermitteln. Daher liefert das Portal keine vollständige Dokumentation der deutschen Rechtsprechung.<sup>6</sup>

Aufgrund der genannten Problematiken ist es sinnvoll weitere mögliche Quellen zu betrachten. Neben Juris kommt noch eine weitere Datenquelle in Frage:

#### Rechtsprechung im Internet

Das Portal *Rechtsprechung-im-Internet*<sup>7</sup> ist eine vom Bundesministerium für Justiz und Verbraucherschutz und dem Bundesamt für Justiz betriebene Webanwendung, auf welcher seit dem Jahr 2010 und teilweise auch rückwirkend öffentlich zugänglich die Gerichtsentscheidungen des Bundesverfassungsgerichtes, der obersten Gerichtshöfe des Bundes<sup>8</sup> sowie des Bundespatentgerichtes veröffentlicht werden. Die Entscheidungen werden dabei ungekürzt und anonymisiert veröffentlicht. Die Datenbank wird laufend aktualisiert und umfasst derzeit etwa 55.000 Gerichtsentscheidungen.<sup>9</sup> Neben einem RSS-Feed, welcher die aktuellsten Gerichtsurteile anzeigt, können einzelne Entscheidungen gezielt über eine filter-basierte Suche aufgerufen und bei Bedarf als PDF und im XML-Format heruntergeladen werden.

Das Portal bietet den Vorteil, dass alle hier zur Verfügung stehenden Entscheidungen frei zur Verfügung stehen und somit die Weiterverarbeitung problemlos möglich ist. Durch die Möglichkeit, Entscheidungen als XML-Datei heruntergeladen zu können, vereinfacht die Plattform die automatisierte Verarbeitung erheblich, da sie eine ein-

---

<sup>5</sup>vgl. [Har20] S. 8

<sup>6</sup>vgl. [Cou18] S. 381

<sup>7</sup>siehe <https://www.rechtsprechung-im-internet.de/>

<sup>8</sup>Bundesgerichtshof, Bundesverwaltungsgericht, Bundesfinanzhof, Bundesarbeitsgericht, sowie Bundessozialgericht

<sup>9</sup>Stand: Dezember 2020

heitlich und grundsätzlich maschinell verwertbare Datenbasis bereitstellt. Als Nachteil ist hier jedoch zu erwähnen, dass *Rechtsprechung-im-Internet* zwar laufend aktualisiert wird, jedoch lediglich ausgewählte Entscheidungen seit dem Jahr 2010 und außerdem nur Entscheidungen der oben genannten Gerichte enthält. *Rechtsprechung-im-Internet* deckt somit ebenfalls nur einen Teil der gesamten Rechtsprechung ab.<sup>10</sup>

Zuletzt lässt sich noch eine weitere Quelle von Entscheidungsdaten betrachten, welche als Grundlage für die Untersuchung in Frage kommt:

### Beck-Online

*beck-online.DIE DATENBANK* (Beck-Online)<sup>11</sup> ist eine juristische Fachdatenbank, welche von dem Verlag C.H. Beck betrieben wird. Innerhalb der Datenbank von Beck-Online werden neben Gerichtsentscheidungen und Fachzeitschriften und auch weitere Literatur und Kommentare veröffentlicht. Mit ungefähr 3,5 Millionen Fällen stellt Beck-Online unter den hier vorgestellten Datenquellen für Entscheidungsdaten die umfangreichste Datenquelle dar<sup>12</sup>. Das Angebot von Beck-Online ist allerdings ebenso wie Juris kostenpflichtig. Bei Beck-Online werden dabei ähnlich zu Juris abonnementbasiert, einzelne Module angeboten, welche sich jedoch im Umfang von den Modulen bei Juris unterscheiden.

Um eine gute Datenauswertung ausführen zu können, sollte ein Datensatz möglichst viele der folgenden fünf Bedingungen erfüllen: Die Daten sollten *frei zugänglich, aus offizieller Quelle, vollständig, unverändert* und *unmittelbar automatisiert weiterverwertbar* sein<sup>13</sup>. Die eben aufgeführten Plattformen zeigen allerdings, dass es zurzeit keine Datenquelle gibt, welche diese Bedingungen vollständig erfüllt und somit als "optimale" Datenquelle infrage käme. Juris und Beck-Online bieten Entscheidungsdaten aus offizieller Quelle, welche unverändert vorliegen<sup>14</sup>, jedoch nicht vollständig und zudem nicht frei zugänglich sind. Über die Weiterverwertbarkeit kann aufgrund des fehlenden Zugangs an dieser Stelle keine Aussage getroffen werden. Eine Auswertung der Daten von Juris oder Beck-Online ist somit für dieses Experiment nicht geeignet. *Rechtsprechung-im-Internet* verfügt ebenfalls nicht über einen vollständigen Datensatz aller Entscheidungen, gestattet dafür jedoch einen freien Zugang zu unveränderten Daten aus offizieller Quelle. Zudem können die Daten effektiv automatisiert erhoben und dann weiterverarbeitet werden, weswegen im Rahmen dieses Experimentes dessen Entscheidungsdaten als Datengrundlage herangezogen werden.

<sup>10</sup>vgl. [Cou18] S. 381

<sup>11</sup>siehe <https://beck-online.beck.de/>

<sup>12</sup>vgl. [Har20] S. 9

<sup>13</sup>vgl. [Cou18] S. 380 und [Har20] S. 14f.

<sup>14</sup>Hierbei sei die Anonymisierung von personenbezogenen Daten ausgenommen.

### 4.1.3 Aufbereitung der Daten

Um die gesammelten Daten für die Auswertungen vorbereiten zu können, ist aus technischer Sicht für die Weiterverarbeitung bedeutend, ob es sich um *strukturierte Daten* oder um *unstrukturierte Daten* handelt. Für die Extraktion der Zitate und die sprachtechnologische Analyse bilden die Entscheidungstexte den Kern der Daten und sind somit für die Auswertung besonders bedeutend. Der Entscheidungstext einer Entscheidung liegt grundsätzlich in textueller Form vor. Bei rein textuellen Daten handelt es sich hier grundsätzlich um unstrukturierte Daten. Zum Vorteil dieser Untersuchung liegt eine Gerichtsentscheidung allerdings in einem XML-basierten Format vor, welches eine einfache Extraktion von Metadaten, wie dem Aktenzeichen, dem Entscheidungsdatum und weiteren Informationen ermöglicht. Somit kann bei diesen Daten auch von *semi-strukturierten* Daten gesprochen werden.

Tabelle 4.1 zeigt, welche Informationen aus einem XML-Dokument einer Gerichtsentscheidung herausgelesen werden können. Beim Einlesen der Dokumente fallen bereits bestimmte Besonderheiten auf. Um die Entscheidungen abspeichern zu können, wird eine ID, also ein eindeutiger Schlüssel benötigt, mit der die Entscheidungen eindeutig voneinander identifizierbar sind. Dies ist vor allem auch für die spätere Konstruktion des Entscheidungsnetzwerkes relevant, da dort ebenfalls jeder Knoten einem eindeutigen Entscheidungsdokument zugeordnet werden soll. Hier würde sich auf den ersten Blick der ECLI einer Entscheidung anbieten. Es besitzt jedoch nicht jede Entscheidung in der Datenbank einen zugeordneten ECLI, da dieser erst 2011 eingeführt wurde. Da auch ältere Entscheidungen im Korpus enthalten sind, scheidet der ECLI aus. Ebenso würde sich anbieten, die Entscheidungen über ihr Aktenzeichen eindeutig zu kennzeichnen. Problematisch ist dabei jedoch, dass ein Aktenzeichen zwar einen juristischen Fall eindeutig bestimmt, nicht jedoch eine einzelne Entscheidung. So können zu einem Dokument mehrere Aktenzeichen oder zu einem Aktenzeichen möglicherweise mehrere Dokumente gehören.<sup>15</sup> Eine eindeutige Zuordnung eines Dokuments über das Aktenzeichen ist somit nur eingeschränkt möglich. Anhand der Metadaten einer Entscheidung würde sich zuletzt auch eine Identifikation über die Dokumentennummer anbieten. Alle Entscheidungen in der *Rechtsprechung-im-Internet* Datenbank haben eine eigene Dokumentennummer, an der sich teilweise sogar bereits die Gerichtsform erkennen lässt. Diese Variante bedarf jedoch der Vorsicht, da die Dokumentennummer zwar eindeutig ist, jedoch nicht offiziell zugeordnet werden kann, da die Zuordnung der Dokumentennummer zu einem Entscheidungsdokument von *Rechtsprechung-Im-Internet* selbst erfolgt und kein klares, allgemeingültiges Benennungsschema dahinter zu erkennen ist.

<sup>15</sup>Dies tritt beispielsweise auf, wenn mehrere Verfahren gemeinsam entschieden werden oder im gleichen Verfahren in verschiedenen Zeitpunkten zunächst verfahrensleitende Beschlüsse und später ein Urteil ergehen und sich beide im Korpus finden.

<b>Dokumentennummer</b>	Eine eindeutige Nummer eines Dokuments.
<b>ECLI</b>	Der European Case Law Identifier (ECLI) bestimmt eine Gerichtsentscheidung anhand eines eindeutigen Identifiers.
<b>Gerichtstyp</b>	Das Gericht, in dem die Entscheidung verabschiedet wurde.
<b>Spruchkörper</b>	Das Organ innerhalb eines Gerichtes, welches die Entscheidung verabschiedet hat.
<b>Entscheidungsdatum</b>	Das Datum, an dem die Entscheidung gefällt wurde.
<b>Aktenzeichen</b>	Das Aktenzeichen des Falles.
<b>Entscheidungstyp</b>	Spezifiziert, ob es sich um ein Urteil oder um einen Beschluss handelt.
<b>Normen</b>	Die Normen, welche in der Entscheidung verwendet wurden.
<b>Vorinstanzen</b>	Alle Vorinstanzen der Entscheidung.
<b>Titel</b>	Der Titel der Entscheidung.
<b>Leitsatz</b>	Der Leitsatz einer Entscheidung.
<b>Sonstige Orientierungssätze</b>	Sonstige Orientierungssätze zu der Entscheidung.
<b>Tenor</b>	Der Tenor der Entscheidung.
<b>Tatbestand</b>	Der Tatbestand einer Entscheidung.
<b>Entscheidungsgründe</b>	Die Gründe für die Entscheidung.
<b>Abweichende Meinungen</b>	Abweichende Meinungen zu der Entscheidung.
<b>URL</b>	Die URL zu der Entscheidung in der Datenbank.

Tabelle 4.1: Metainformationen eines Entscheidungsdokumentes

Würde die Datenbasis um Entscheidungen anderer Datenquellen ergänzt, ließen sich dort keine eindeutigen Identifier mit diesem System erzeugen.

Um die Entscheidungen innerhalb dieses Korpus eindeutig zuordnen zu können, wird das Aktenzeichen verwendet. In dieser Arbeit soll evaluiert werden, ob sich ein Zusammenhang von thematisch verwandten Entscheidungen über ein Zitatnetzwerk abbilden lässt. Teilweise werden in einem gerichtlichen Urteil oder Beschluss mehrere Verfahren mit mehreren Aktenzeichen zusammen entschieden. Diesen Verfahren werden dann jeweils diese Aktenzeichen zugeordnet. Ein Aktenzeichen führt also immer zu einer Entscheidung, aber eine Entscheidung kann mehrere Aktenzeichen haben. Da diese den gleichen Gegenstand bezeichnen, können sie synonym verwendet werden.

Zusätzlich ist zu erwähnen, dass auch die Zuordnung der oben genannten Metadaten

innerhalb eines XML-Entscheidungsdokuments nicht überall gleich erfolgt. So liegen die Entscheidungsgründe einer Entscheidung teilweise als <gruende> und teilweise als <entscheidungsgruende> vor. Hier sind beim Einlesen der Daten somit zusätzliche Überprüfungen notwendig. Außerdem sind fast alle Elemente des Entscheidungstextes<sup>16</sup>, bis auf den Tenor innerhalb eines Entscheidungstextes mit einer Randnummer<sup>17</sup> versehen, welche zusätzlich zum dazugehörigen Text abgespeichert werden muss.

Um die Informationen für Analysen also strukturiert, vollständig und gut abrufbar zu speichern, wird im Rahmen dieser Arbeit für jede Gerichtsentscheidung ein Entscheidungsobjekt **Decision** erstellt, in welchem die Informationen abgelegt werden.<sup>18</sup> Um ein Zitationsnetzwerk aus Gerichtsentscheidungen erzeugen zu können, ist es weiterhin notwendig, die auftretenden Zitationen in einem Dokument zu extrahieren und so auflösen zu können, dass sie einem anderen Dokument zugeordnet können und eine entsprechende Kante gezeichnet werden kann. Diese Aufgabe wirkt zunächst trivial, da intuitiv davon ausgegangen wird, dass eine Entscheidung – welche sich eindeutig durch das zugehörige Aktenzeichen identifiziert – direkt über diesen Identifier auf eine andere Entscheidung deutet. In Gerichtsentscheidungen werden jedoch in den meisten Fällen nicht direkt die Aktenzeichen eines Urteils zitiert, sondern eine zum Aktenzeichen gehörige *Fundstelle*.<sup>19</sup>

Die Verwendung von Fundstellen statt Aktenzeichen kann historisch erklärt werden. Da früher keine digitalen Datenbanken existierten, in welchen Urteile und Beschlüsse unter freiem Zugang nachgeschlagen werden konnten, war es Juristen nicht ohne weiteres möglich, zu erkennen, wo ein zitiertes Urteil einzusehen ist. Hierfür wurden Gerichtsurteile meist in einer juristischen Fachzeitschrift veröffentlicht und dann nach diesem einfach aufzufindenden Erscheinungsort zitiert. Andernfalls hätten Rechtswissenschaftler jeweils einzeln mit dem Aktenzeichen bei den Gerichten eine Kopie beantragen müssen, was nicht praktikabel erschien. Alternativ wurden für besonders bedeutend gehaltene Gerichtsentscheidungen zu sog. *amtlichen Sammlungen*<sup>20</sup> hinzugefügt, wodurch sie ebenfalls eine Fundstelle erhielten. Eine Fundstelle liegt dabei zumeist in folgendem Format vor:

{Sammlung | Zeitschrift} {Jahr},{Seite}

Die Fundstelle beginnt mit dem Kürzel der amtlichen Sammlung oder dem Kürzel der Zeitschrift, worauf die Angabe des Jahres der Veröffentlichung der Zeitschrift und die

<sup>16</sup>Dies bezieht sich auf den Tatbestand, die Entscheidungsgründe und die abweichenden Meinungen.

<sup>17</sup>Die Randnummern werden dazu verwendet, um exakt auf gewisse Stellen in Entscheidungstexten verweisen zu können.

<sup>18</sup>siehe Abschnitt 4.1.4

<sup>19</sup>Eine Ausnahme bildet hier der BGH, welcher Entscheidungen teilweise mit Aktenzeichen und Fundstelle zitiert. Beispiel, siehe: BGH, Urteil vom 04.11.2020 - IV ZR 19/19 Rn. 16

<sup>20</sup>Liste aller bedeutenden Entscheidungssammlungen: <https://t1p.de/jw0q>

Angabe der Seite, an der die Gerichtsentscheidung beginnt, erfolgt. Beispiele für diese Notation sind: *NJW 2020, 384* und *BVerfGE 69, 141*. In der Praxis gibt es allerdings verschiedene Sonderformen, welche eine vollständige Extraktion aller Zitate verkomplizieren. So werden z.B. in der amtlichen Sammlung des Bundesverfassungsgerichtes Entscheidungen teilweise in der oben genannten Schreibweise zitiert, teilweise auch als sog. *Langfundstelle*<sup>21</sup>, bei der neben der Startseite auch die Endseite des Urteils angegeben wird. Eine Angabe als Langfundstelle erfolgt am Beispiel des Bundesverfassungsgerichtes im folgenden Format:

BVerfGE {Jahr},{Startseite}-{Endseite}

Ob eine Entscheidung dabei in einer Zeitschrift veröffentlicht wird, obliegt neben dem Richter bzw. Spruchkörper auch dem Verlag der Zeitschrift. Ähnlich ist es bei den Entscheidungssammlungen. Hier entscheiden die Richter, die ein Urteil verabschiedet haben, selbst, ob die Entscheidung Teil der Sammlung werden soll oder nicht.

Die unterschiedlichen Verfahren führen dazu, dass eine Entscheidung somit sowohl eine Fundstelle aus einer Veröffentlichung in einer Fachzeitschrift, als auch eine Fundstelle aus einer amtlichen Entscheidungssammlung enthalten kann. Ein Beispiel bildet die Entscheidung des Bundesverfassungsgerichtes vom 27.02.2018 - 2 BvE 1/16. Auf diese Entscheidung kann über die Fundstellen *BVerfGE 148, 11*; *NJW 2018, 928*; *NVwZ 2018, 485*; *DVBl 2018, 503*; *KR 2018, 245* oder *DÖV 2018, 374* verwiesen werden. Eine Entscheidung kann somit also über verschiedene Fundstellen gleichermaßen zitiert werden.

Um ein Netzwerk über die Zitate der Entscheidungen realisieren zu können, muss eine Struktur erzeugt werden, die eine Fundstelle auf das Aktenzeichen einer Entscheidung abbildet (*Mapping*). Diese Informationen lassen sich aus den bereits gewonnenen Daten nicht ableiten, sodass weitere Datenquellen hinzugezogen werden müssen. Derzeit gibt es keine Datenquelle, die solche Informationen frei zugänglich, offiziell, vollständig, unverändert und automatisiert weiterverwendbar anbietet. Auf der Webseite des Bundesverfassungsgerichtes<sup>22</sup> kann z.B. die amtliche Sammlung BVerfGE mit den dazugehörigen Entscheidungen eingesehen werden. Eine solche Quelle eignet sich für eine automatisierte Weiterverarbeitung, enthält jedoch nur ausgewählte Entscheidungen des Bundesverfassungsgerichtes. Die oben bereits angeführte Quelle *Juris*<sup>23</sup> bietet ebenfalls Daten, mit denen sich Fundstellen auf Entscheidungen abbilden lassen, zwar an, jedoch nicht vollständig und auch nicht frei zugänglich. An den eben genannten Beispielen wird deutlich, dass das Erstellen einer Mapping-Struktur nur auf Basis von

<sup>21</sup>vgl. [Cou19] S. 233

<sup>22</sup>Siehe [https://www.bundesverfassungsgericht.de/DE/Entscheidungen/Entscheidungen/Amtliche\\_Sammlung\\_BVerfGE.html](https://www.bundesverfassungsgericht.de/DE/Entscheidungen/Entscheidungen/Amtliche_Sammlung_BVerfGE.html)

<sup>23</sup>siehe 4.1.2

#### 4 Daten

verschiedenen Quellen erfolgen kann, wodurch eine gute Datenqualität im Bezug auf Vollständigkeit, freie Zugänglichkeit, Unverändertheit, Weiterverwertbarkeit und offizieller Zugehörigkeit nur schwer zu erreichen ist. Da die verschiedenen Datenquellen auch unterschiedlich weiterverarbeitet werden, entsteht hier zusätzlich ein großer Zeitaufwand, um die Daten zu sammeln und aufzubereiten.

Eine Möglichkeit zumindest an einen Teil solcher Daten zu gelangen, bietet das oben bereits genannte Web-Portal Beck-Online. Beck-Online ist aufgrund der nötigen Zahlungspflicht ungeeignet als Quelle, um Daten von Gerichtsentscheidungen zu gewinnen, bietet allerdings die Möglichkeit juristische Fachzeitschriften und im genaueren den Aufbau ihrer Inhalte kostenfrei anzusehen. Innerhalb des Inhaltsverzeichnisses eines Heftes sind unter dem Abschnitt *Rechtsprechung*<sup>24</sup> alle Entscheidungen dieses Heftes mit Nennung ihrer Aktenzeichens aufgelistet, welche in Zusammenhang mit der Seitenzahl, an der die Entscheidung beginnt, zu ihrer Fundstelle zusammengesetzt werden können. Das Format auf Beck-Online ist in Tabelle 4.2 an einem Beispiel der *Neuen Juristischen Wochenschrift* (NJW) illustriert.<sup>25</sup>

#### Inhaltsverzeichnis

NJW Heft 47/2020

[...]

#### Rechtsprechung

[...]

StGH Hess, 12.08.2020 - P.St. 2689: Erschöpfung des Rechtswegs in Erbscheinsverfahren 3436

BGH, 27.07.2020 - VI ZR 405/18: \*Entfernung von Links aus Suchergebnislisten 3436

BGH, 27.07.2020 - VI ZR 476/18: Folgen von Rechtsschutz und Kontext bei Auslistungsbegehren 3444

[...]

BGH, 27.02.2020 - III ZB 61/19: Befangenheitsablehnung im Berufungsverfahren 3460

[...]

Tabelle 4.2: Auszug aus dem Inhaltsverzeichnis der NJW 2020, Heft 47

Anhand des Namens der Zeitschrift, des Jahres und der Seite, auf der die Entscheidung beginnt, lässt sich nun aus diesen Daten die Fundstelle einer Entscheidung generieren, welche auf das im Text angegebene Aktenzeichen abgebildet werden kann. Aus dem Ausschnitt lassen sich, basierend auf dem eben vorgestellten Schema, die folgenden Abbildungen von Fundstellen auf Aktenzeichen bestimmen: Die Fundstelle *NJW 2020*,

<sup>24</sup>In einigen Heften wird die Rechtsprechung auch unter den Abschnitten "Entscheidungen" oder "Mitteilungen" aufgeführt. Der Abschnitt "Rechtsprechung" tritt jedoch am häufigsten auf.

<sup>25</sup>Auszug aus NJW 2020, Heft 47, siehe <https://t1p.de/qtsv>

3436 verweist auf die Entscheidung *VI ZR 2020/18*, Fundstelle *NJW 2020, 3444* verweist auf die Entscheidung *VI ZR 476/18* und die Entscheidung *III ZB 61/19* kann über die Fundstelle *NJW 2020, 3460* gefunden werden. So können, aufgebaut nach Jahr der Veröffentlichung, die einzelnen Hefte einer Zeitschrift mit ihrem Inhaltsverzeichnis einschließlich der Seitenzahlen eingesehen werden, ohne einen Dienst zahlungspflichtig zu erwerben. Um diese Informationen automatisiert abrufen zu können, lassen sich die Parameter der Webseite-URL anpassen. Der Aufruf eines Heftes verfolgt das folgende Schema:

```
https://beck-online.beck.de/?vpath=bibdata/zeits/<Zeitschriftenkürzel>  
/2020/cont/NJW.<Jahr>.H<HeftNr>HALTSVERZEICHNIS.htm
```

Hier ist im Vorfeld jedoch Einiges zu beachten: Es wird nicht jedes Jahr die gleiche Anzahl an Heften veröffentlicht, wodurch dies vor der Abfrage in Erfahrung gebracht werden muss. Des Weiteren werden in jedem Jahr teilweise verschiedene Hefte zu einem Heft zusammengefasst, wodurch in diesen Fällen die Heftnummer in der URL nicht nur aus einer, sondern beiden Heftnummern besteht. Da dies in jedem Heft beliebig passiert, müssen solche Fälle ebenfalls vorher untersucht und aufgefangen werden. Für diese Untersuchung wurden alle Hefte überprüft, um die Anzahl der Hefte sowie die zusammengefassten Hefte zu ermitteln.

Das Verfahren am Beispiel der Neuen Juristischen Wochenschrift (NJW) anhand des Beck-Online Portals lässt bereits den Aufwand für eine vollständige Struktur zur Abbildung aller Fundstellen erahnen. Es müssten alle Entscheidungssammlungen und alle juristischen Fachzeitschriften automatisiert durchsucht und die Fundstellen extrahiert werden, was eine große Anzahl verschiedener Quellen und verschiedener Verarbeitungsverfahren bedeuten würde, da Beck-Online nicht alle juristischen Fachzeitschriften in seiner Datenbank führt und die verschiedenen Fachzeitschriften zudem unterschiedlich aufgebaut sind und somit das Extrahieren der Fundstellen nicht genau so wie bei der NJW erfolgen kann. Es wird deutlich, dass eine Abbildung aller Zitate den Rahmen dieser Arbeit übersteigt.

Im Rahmen der Auswertungen dieser Veröffentlichung werden zunächst nur die Zitate betrachtet, dessen Fundstellen aus der NJW der Jahre 1948 bis 2020<sup>26</sup> sowie aller Entscheidungen der amtlichen Sammlung des Bundesverfassungsgerichtes<sup>27</sup> stammen. Diese Auswahl liegt darin begründet, dass die NJW als zentrales juristisches Informationsmedium von solchem Rang gilt, dass vom Bundesgerichtshof unter anderem verlangt wird, dass jeder Rechtsanwalt und jede Rechtsanwältin die Inhalte der NJW regelmäßig zu Kenntnis nimmt und bei Haftungsfragen alle Urteile, die in dieser Zeit-

<sup>26</sup>Im Jahr 2020 wurden bis zum Zeitpunkt der Auswertungen die Daten bis Heft 45 Daten gesammelt.

<sup>27</sup>Für die Auswertung werden die Daten bis Band 152 berücksichtigt.

schrift erschienen sind, als bekannt vorausgesetzt werden können.<sup>28</sup>

Mit den Fundstellendaten sind nun alle notwendigen Daten vorhanden, um anschließend Analysen darauf ausführen zu können. Dieser Arbeit liegt jedoch neben der Datenanalyse noch ein weiterer Schwerpunkt zugrunde: Das Erstellen einer Softwareinfrastruktur, welche die Entscheidungen für eine Analyse vorbereitet.

### 4.1.4 Software

Auf der Basis der beschriebenen Datengrundlage soll nun ein System entwickelt werden, welches die Gerichtsentscheidungen persistent speichert, auf einem aktuellen Stand hält und somit eine Infrastruktur bietet, um Analysen auf die gegenwärtigen Daten auszuführen. Der folgende Abschnitt wird ein auf Basis der Programmiersprache *Java* entwickeltes System erläutern, welches diese Anforderungen aufgreift und technisch löst. Hierfür werden verschiedene Technologien verwendet, welche in diesem Kapitel zu Beginn vorgestellt werden. Darüber hinaus wird begründet, weshalb gerade diese verwendeten Technologien eingesetzt werden. Es wird außerdem auf die Architektur der Komponenten des Systems eingegangen, und zuletzt der Datenfluss eines Dokuments im System veranschaulicht, um verdeutlichen, welche Stationen ein Dokument innerhalb des Systems durchläuft.

#### Tools

Im hier entwickelten Softwaresystem kommen die folgenden Technologien zum Einsatz:

#### Elasticsearch

*Elasticsearch*<sup>29</sup> ist eine verteilte Suchmaschine, welche ihre Daten als JavaScript-Object-Notation (JSON) Dokumente im NoSQL Format abspeichert. Elasticsearch basiert auf der Volltextsuche von Apache Lucene, ist gut skalierbar und sehr effizient, wodurch es sich sehr gut für die Verarbeitung großer Datenmengen eignet. Elasticsearch verwendet für die Indexierung und die Suchanfragen eine RESTful<sup>30</sup>-API. Es ist Teil des Elastic Stacks und somit einfach mit Anwendungen wie Kibana, Beats und Logstash kombinierbar. Der Elastic Stack steht als *Open-Core-Modell*, somit also teilweise Open-Source zur Verfügung. Elasticsearch wird in diesem Experiment als Datenbank zur persistenten Speicherung der Entscheidungsdaten verwendet. Elasticsearch wurde aufgrund der guten Skalierbarkeit, der hohen Effizienz und der einfachen Integration für dieses System ausgewählt.

<sup>28</sup>vgl. BGH, Beschluß vom 20. 12. 1978 - IV ZB 115/78, NJW 1979, 877

<sup>29</sup><https://www.elastic.co/de/elasticsearch/>

<sup>30</sup>REST = Representational State Transfer

**Kibana**

*Kibana*<sup>31</sup> fungiert als webbasierte Benutzeroberfläche für Elasticsearch und bietet die Möglichkeit, Elasticsearch-Daten filterbasiert zu suchen und zu visualisieren. Kibana ist ebenfalls Teil des Elastic Stack und somit Open-Source. Kibana wurde für dieses System aufgrund der Zugehörigkeit zu Elasticsearch, der schnellen Integration und der funktionalen Möglichkeiten herangezogen.

**Spring Framework**

Das Framework *Spring*<sup>32</sup> bietet verschiedene Unterstützungsmöglichkeiten, um mit geringem Aufwand größere Java-Anwendungen zu realisieren. Das Framework kann dabei durch modulare Erweiterungen wie z.B. *Spring Web*, *Spring Data* ergänzt werden, um komplexere Web-Anwendungen zu erstellen oder Datenbank Abfragen einfach realisieren zu können. Neben Spring Web, welches für die Entwicklung der Webseite verwendet wurde, ermöglicht das Modul Spring Data, innerhalb dieses Systems mit wenigen Zeilen Code Anfragen an den Elasticsearch Server zu stellen. Zusätzlich ermöglicht das Modul *Spring Boot*, ein schnelles Starten des Systems sowie eine stetige Erweiterung der Datenbasis um aktuell veröffentlichte Entscheidungen. Aufgrund der vielseitigen modularen Erweiterungsmöglichkeiten, die einfache Integration in das System und die dadurch erreichte Ersparnis an Implementationsaufwand wurde das Spring Framework für dieses System ausgewählt.

**OpenNLP**

*OpenNLP*<sup>33</sup> ist eine von der *Apache Software Foundation* zur Verfügung gestellte Bibliothek für die Programmiersprache Java, welche Werkzeuge des maschinellen Lernens für NLP-Aufgaben bereitstellt. Mit diesen Werkzeugen können typische Aufgaben aus der natürlichen Sprachverarbeitung wie z.B. Named Entity Extraction, Satzsegmentierung, Tokenisierung oder die Identifikation einer Sprache leicht umsetzbar implementiert werden. OpenNLP unterliegt einer Open-Source Lizenz und wird stetig weiterentwickelt, weswegen es für diese Untersuchung ausgewählt wurde. Innerhalb dieses Projektes wurde OpenNLP für die Tokenisierung und die Satzsegmentierung von Elementen aus den Entscheidungstexten verwendet.

**Jsoup**

Das Framework *Jsoup*<sup>34</sup> bietet die Möglichkeit, automatisiert Webseiten aufzurufen. Jsoup ist Open-Source und bietet dem Nutzer zudem die Funktionalität für *Web Scra-*

---

<sup>31</sup><https://www.elastic.co/de/kibana>

<sup>32</sup><https://spring.io/projects/spring-framework>

<sup>33</sup><https://opennlp.apache.org/>

<sup>34</sup><https://jsoup.org/>

#### 4 Daten

*ping*<sup>35</sup>, um automatisiert Informationen von Webseiten zu erhalten. Jsoup ruft dabei die Webseite auf und parst ihren Quelltext, sodass anschließend Informationen daraus gewonnen werden können. Die Funktionalität von Jsoup wurde beim Gewinnen der Entscheidungsdokumente von der Seite *Rechtsprechung-im-Internet* verwendet. Gegenüber anderen Frameworks bietet Jsoup die Option mit geringem Aufwand schnell Daten im XML- und HTML Format zu parsen und weiter zu verarbeiten, weswegen sich in der Anwendung dieser Arbeit für dieses Framework entschieden wurde.

#### Selenium

Das Framework *Selenium*<sup>36</sup> bietet – ähnlich wie Jsoup – die Möglichkeit, automatisiert Webanwendungen aufzurufen und zu testen. Anders als bei Jsoup simuliert das System allerdings einen Nutzer, welcher über einen Browser die Webseite besucht. Selenium ermöglicht somit auch das Aufrufen von Inhalten, welche z.B. mit JavaScript geladen werden. Innerhalb dieses Systems wurde Selenium für die Gewinnung der Fundstellendaten von Beck-Online verwendet. Da diese mit JavaScript generiert werden, konnte das Jsoup Framework nicht verwendet werden, weswegen hierfür Selenium herangezogen wurde. Auch Selenium ist Open-Source und wurde aufgrund der schnellen Integration und der leichten Handhabung im System dieser Arbeit verwendet.

#### JGraphT

*JGraphT*<sup>37</sup> ist eine Bibliothek für die Programmiersprache Java, welche Datenstrukturen und Algorithmen der Graphentheorie bereitstellt. Seit diesem Jahr stellt JGraphT zusätzlich eine Instanz für die Programmiersprache Python bereit. JGraphT enthält sehr effiziente und generische Datenstrukturen für Graphen und stellt außerdem eine große Anzahl an verschiedenen State of the Art-Algorithmen bereit, weswegen sich in diesem Projekt für diese Bibliothek entschieden wurde.

#### Systemkomponenten

In Zusammenhang mit den oben genannten Tools werden nun die Komponenten der Software beschrieben. Es wird mit der Eingliederung der Daten in das Softwaresystem begonnen.

---

<sup>35</sup>Der Begriff *Web Scraping* bezeichnet das automatisierte Extrahieren von Informationen einer Webseite, um von dieser Informationen zu gewinnen.

<sup>36</sup><https://www.selenium.dev/>

<sup>37</sup><https://jgrapht.org/>

Die Software startet initial damit, dass zuerst die Daten über die im Vorfeld ausgesuchte Quelle *Rechtsprechung-im-Internet* heruntergeladen werden. Hier wird in zwei Schritten vorgegangen: Zuerst wird eine Möglichkeit gesucht, rückwirkend alle bereits in der Datenbasis des Web-Portals vorhandenen Entscheidungen herunterzuladen. Im zweiten Schritt wird am Portal eine Schnittstelle gesucht, um zu erkennen, ob es neu veröffentlichte Entscheidungen gibt und diese anschließend ebenfalls herunterzuladen.

Für den ersten Schritt bestehen auf der Webseite verschiedene Möglichkeiten, um die Entscheidungsdaten herunterzuladen. Grundsätzlich können über einen emulierten automatisierten Webbrowser (so wie Selenium ihn beispielsweise anbietet) die Entscheidungen in einer Liste automatisiert aufgerufen und schließlich auf der Detailseite einer Entscheidung das XML Dokument heruntergeladen werden. Die Emulation eines Browsers ist hier zwingend notwendig, da die Liste der Entscheidungen im Portal per Javascript befüllt wird. *Rechtsprechung-im-Internet* bietet allerdings noch eine weitere Möglichkeit, welche es ermöglicht, die Entscheidungen noch einfacher herunterzuladen. Das Portal bietet auf seiner Webseite eine Liste aller Entscheidungen im XML-Format an, welche über den Aufruf einer URL eingesehen werden kann. Aus dieser lassen sich unter anderem der Download-Link zu jeder Entscheidung und der Zeitpunkt der letzten Modifikation dieser Entscheidung herauslesen. Aus dieser Liste können also alle Download-Links zu den Entscheidungen einfach herausgelesen werden. Da diese Liste nur einmalig aufgerufen werden muss und anschließend lediglich die URL für den Download einer Entscheidung aufgerufen wird, minimiert das gegenüber der ersten Möglichkeit stark den Datenverkehr, da in erstgenannter Möglichkeit pro Entscheidung mehrere Aufrufe notwendig sind. Über diese Möglichkeit wird nun für jede Entscheidung eine ZIP-Datei heruntergeladen, welche das erforderliche XML-Dokument mit den Entscheidungsinformationen enthält.

Schließlich wird eine Möglichkeit gesucht, um neue Entscheidungen automatisiert herunterzuladen zu können. Die dafür zuerst am einfachsten erscheinende Möglichkeit ist der Download über einen vom Portal angebotenen *Rich Site Summary*-Feed (RSS-Feed), welcher die neuesten veröffentlichten Entscheidungen des Portals anzeigt. Im Verlauf der Realisierung dieser Arbeit stellte sich diese Möglichkeit jedoch als ungeeignet heraus, da hier nur neue kürzlich entschiedene Entscheidungen veröffentlicht werden, jedoch keine älteren Entscheidungen, welche rückwirkend in die Datenbasis des Portals aufgenommen werden. Hier wurde schließlich erneut die XML-Liste aus dem ersten Schritt verwendet, da hier über den Zeitpunkt der letzten Modifikation sofort einsehbar ist, welche Entscheidungen insgesamt neu in der Datenbasis des Portals sind. Um die neuen Entscheidungen in die Datenbasis dieses Projektes fortlaufend zu integrieren, fragt die Software mit einem **Scheduling Job** täglich das Portal ab und lädt, wenn notwendig, die neuen Entscheidungen herunter. Hier kommt das Spring

Framework, mit der Komponente Spring Boot zum Einsatz. Die Abfrage geschieht täglich um 24 Uhr. Dies hat zum einen den Grund, den Datenverkehr möglichst auf eine Uhrzeit zu legen, an der das Portal wenig besucht wird. Zum anderen werden die neuen Entscheidungen bei *Rechtsprechung-im-Internet* ausschließlich zwischen 22 Uhr und 24 Uhr veröffentlicht, weswegen die Datenbasis dieses Projektes in kürzester Zeit auf dem aktuellsten Stand ist.

Aus dem heruntergeladenen XML-Dokument werden nun die Daten der Entscheidungen extrahiert. Dies wird in der **Vorverarbeitung** von einem *DecisionMapper* innerhalb der Software durchgeführt und dient dem Zweck, die Daten in ein Format zu bringen, welches für weitere Analysen leicht weiter verwertbar ist. Dazu gehören alle in Tabelle 4.1 bereits aufgeführten Informationen. Zusätzlich zu den Daten des XML werden außerdem die im Text auftretenden Fundstellen anderer Entscheidungen aus Titel, Leitsatz, den sonstigen Orientierungssätzen, Tenor, Tatbestand, Entscheidungsgründen und abweichenden Meinungen extrahiert. Um die gefundenen Fundstellen in das mit ihnen verbundene Aktenzeichen zu übersetzen, wird anschließend das Mapping mit den Daten von Beck-Online durchgeführt, sodass nun erkennbar ist, welches Aktenzeichen genau zitiert wird. Diese Daten werden für die Erzeugung des Entscheidungsnetzwerkes benötigt. Sind alle nötigen Informationen für eine Entscheidung extrahiert, werden

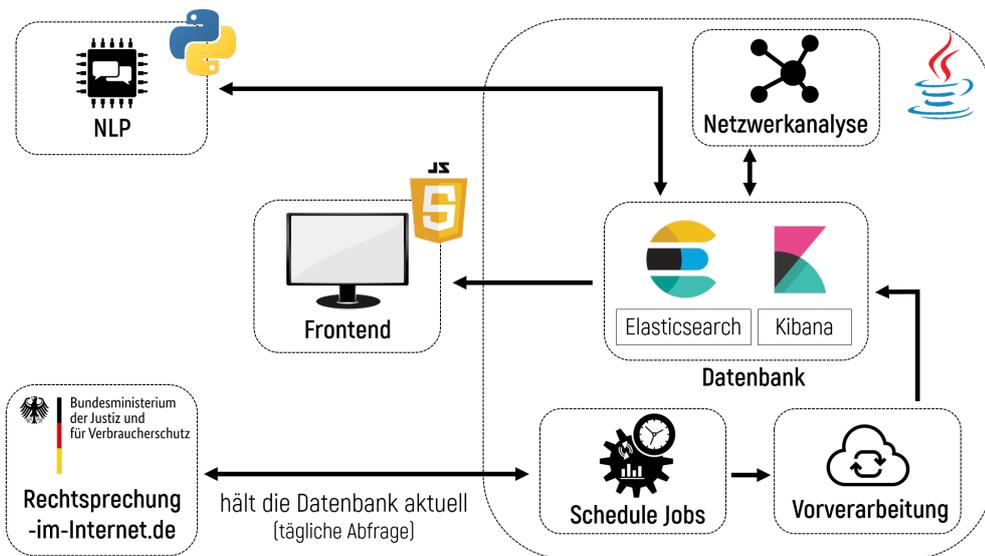


Abbildung 4.1: Die Systemarchitektur der Software

diese schließlich in der Elasticsearch **Datenbank** gespeichert. Durch die Verwendung der Komponente Spring Data aus dem Spring Framework kann hier das Hochladen sowie das Abfragen der Daten aus der Datenbank unkompliziert und schnell realisiert werden.

Die Abbildung 4.1 stellt die Architektur der Software und die Zusammenhänge der Komponenten dar. Das Konstrukt auf der rechten Seite visualisiert hier die Java Anwendung, welche die Entscheidungen sammelt, verarbeitet und in der Datenbank speichert. Diese bezieht die neuen Entscheidungen von dem Webportal *Rechtsprechung-im-Internet*, wobei die Anfrage täglich über den Schedule Job realisiert wird.

Auf Basis der Daten, die nun in der Elasticsearch Datenbank vorhanden sind, wird dann mithilfe von JGraphT das Entscheidungsnetzwerk generiert und die **Netzwerk-analyse** durchgeführt. Dies passiert regelmäßig in einem Abstand von sieben Tagen, automatisiert über den Schedule Job. Anhand verschiedener Berechnungen, welche auf den Graphen ausgeführt werden, ergeben sich schließlich weitere Informationen, welche anschließend in der Datenbank abgespeichert werden.<sup>38</sup> Da sich die Datenbasis stetig erweitert, wird das Entscheidungsnetzwerk in regelmäßigen Intervallen neu generiert. Dafür ist ebenfalls der bereits vorgestellte Scheduling Job zuständig.

Die Anwendung ruft außerdem auch das **NLP** Aufgaben, also das Klassifikationsmodell auf. Um die maschinellen Lernalgorithmen umzusetzen, wurde in dieser Arbeit die Programmiersprache *Python* mit den Bibliotheken *Pytorch*<sup>39</sup> und *Keras*<sup>40</sup> verwendet. Um die sprachtechnologischen Aufgaben mit der Java Applikation zu verbinden, wurde hierfür ein Service entwickelt, mit dem der Python Code aus der Java Anwendung aufrufbar ist.

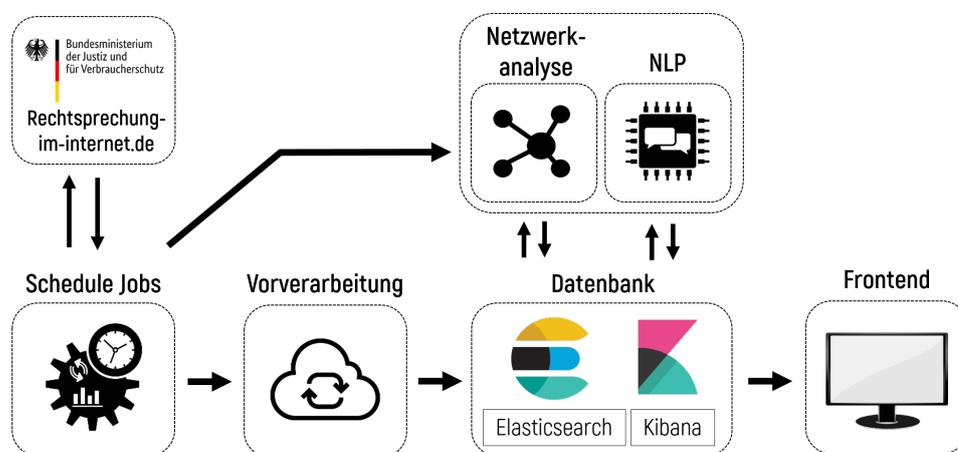


Abbildung 4.2: Der Datenfluss innerhalb der Architektur

In Abbildung 4.2 wird der Datenfluss eines Dokumentes durch die Softwarekomponen-

<sup>38</sup>siehe Kapitel 5

<sup>39</sup><https://pytorch.org/>

<sup>40</sup><https://keras.io/>

ten dargestellt. Der Fluss wird dabei durch den Schedule Job initiiert, welcher den Download an der Datenquelle anstößt. Die Entscheidungsinformationen werden nach dem Download schließlich in der Vorverarbeitung für die Speicherung in der Datenbank aufbereitet und anschließend in der Elastic Datenbank gespeichert. Von hier aus können die Daten an das Frontend gesendet werden. Beginnt der Schedule Job eine Netzwerkanalyse oder einen NLP Task, werden die Daten aus der Datenbank abgefragt, wobei die Ergebnisse der Berechnungen wieder in der Datenbank abgespeichert werden.

Um die Ergebnisse dieser Arbeit greifbar zu machen und die Option zu bieten, diese interaktiv zu betrachten, wurde im Rahmen der Untersuchung außerdem ein **Frontend** in Form einer Webanwendung erstellt. Die Webanwendung bietet die Möglichkeit, den Entscheidungsdatensatz in einer Suchmaschine zu durchsuchen und somit einzelne Entscheidungen z.B. per Aktenzeichen oder per Keyword aufzurufen. Es werden außerdem einige Ergebnisse der Untersuchung grafisch dargestellt und es können die generierten Subnetzwerke des Entscheidungsnetzwerkes interaktiv durchsucht werden. Die Weban-

Titel
Wegezeiten - Mitbestimmung des Betriebsrats
Aktenzeichen
1 ABR 11/18
ECLI
ECLI:DE:BAG:2019:221019.B.1ABR11.18.0
Gerichtstyp
BAG
Spruchkörper
1. Senat
Entscheidungsdatum
22.10.2019
Entscheidungstyp
Beschluss
Normen
§ 87 Abs 1 Nr 2 BetrVG
Vorinstanzen

Abbildung 4.3: Die Detailseite einer Entscheidung

wendung ist dabei mit der Datenbasis eng verknüpft und enthält somit zu jeder Zeit die neuesten Entscheidungen und somit die neuesten generierten Netzwerke der Datenbasis. Abbildung 4.3 zeigt die Detailseite einer Entscheidung im Frontend. Am oberen Rand der Abbildung ist zudem die Suche zu sehen, mit welcher nach Entscheidungen gesucht werden kann.

#### 4.1.5 Statistiken des Korpus

Im Rahmen der Datenerhebung für diese Untersuchung wurde nun ein Datensatz zusammengestellt, welcher als Rohstoff für die kommenden Auswertungen dienen soll. Um einen Eindruck erhalten zu können, wie der Datensatz aussieht und aufgebaut ist, werden im folgenden Abschnitt einige Statistiken aufgeführt, welche den Datensatz beschreiben. Da sich der Datensatz automatisch aktuell hält, kann hier allerdings nur eine Momentaufnahme des Datensatzes zum Zeitpunkt der Untersuchung erfolgen.<sup>41</sup> Dies sollte beim Lesen dieses Abschnittes berücksichtigt werden.

Die Tabelle 4.3 demonstriert die Anzahl der Entscheidungen in der Datenbank pro Gericht. Die Anteile der einzelnen Gerichte am Korpus werden in der Tabelle eben-

Gerichtstyp	Anzahl der Entscheidungen	Anteil (in %)
BGH	19.214	35
BFH	9.008	16
BVerwG	7.289	13
BPatG	5.925	11
BAG	5.703	10
BSG	4.509	8
BVerfG	3.930	7
Gesamt	55.580	100

Tabelle 4.3: Anzahl der Entscheidungen im Korpus je Gericht

falls dargestellt. Hier ist deutlich zu erkennen, dass der Bundesgerichtshof mit 35% den größten und das Bundesverfassungsgericht mit 7% den geringsten Anteil der Entscheidungen ausmacht. Um einen Eindruck davon zu erhalten, wie die Datenbasis in Zukunft wachsen wird, können anhand der Betrachtung der Monate genau eines rückwirkenden Jahres grobe Schätzungen vorgenommen werden. Abbildung 4.4 visualisiert die Anzahl der neu zur Datenbasis hinzugekommenen Entscheidungen von Dezember 2019 bis November 2020. Innerhalb des rückwirkenden Jahres sind also insgesamt 3.480 Entscheidungen zur Datenbasis hinzugekommen. Anhand dieser Werte lässt sich berechnen, dass zukünftig im Durchschnitt etwa 10 Entscheidungen pro Tag neu zur Datenbasis dazukommen werden.<sup>42</sup> Bei der Betrachtung der Daten ist zu beobachten, dass unter der Woche zumeist mehr Entscheidungen in die Datenbasis hinzukommen, und am Wochenende weniger.

<sup>41</sup>Stand: Dezember 2020

<sup>42</sup>Bei dem eben genannten Durchschnittswert wurde kaufmännisch gerundet.

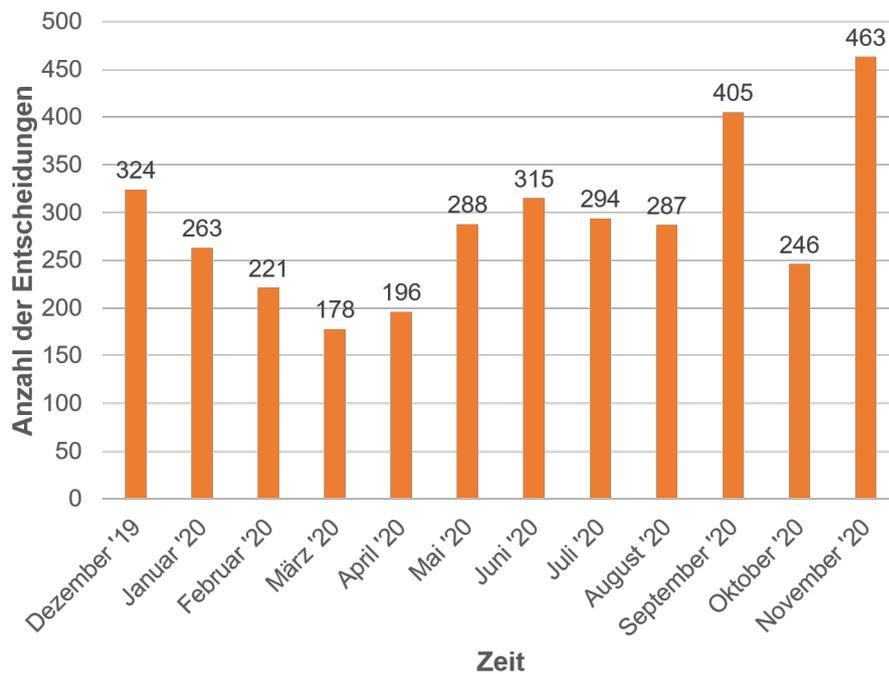


Abbildung 4.4: Anzahl hinzugefügter Entscheidungen je Monat

Um auf diesen Datensatz nun Analysen ausführen zu können, muss nun schließlich die Methodik dafür evaluiert werden. Da innerhalb dieser Arbeit auch methodische Ansätze des maschinellen Lernens zum Einsatz kommen, werden allerdings vorher noch Trainings- und Testdaten benötigt.

## 4.2 Trainingsdaten und Testdaten

Um die methodischen Ansätze dieser Arbeit evaluieren zu können, werden dafür Daten benötigt, welche aus einer objektiven Betrachtung möglichst "korrekt" sind. Um solche Daten zu erhalten, ist es sinnvoll die Rohdaten, welche hier in textueller Form vorliegen, einem Experten vorzulegen, welcher diese Daten annotiert, somit also die erforderlichen Stellen im Text einer Klasse zuordnet. Diese Daten werden dann als sog. *Grundwahrheit* (engl. *Ground Truth*) bezeichnet. Die Ergebnisse des eigenen Ansatzes werden diesen Daten schließlich gegenübergestellt. Anhand des Vergleiches der eigenen Methodik mit diesen Daten, kann die Güte des eigenen Ansatzes messbar gemacht werden.

Aus dem eben vorgestellten Korpus wurden hierfür zufällig insgesamt 140 Entscheidungs-

gen für einen Trainings- und Testdatensatz ausgewählt. Dabei wurde berücksichtigt, dass insgesamt 20 Entscheidungen von jeder Gerichtsbarkeit Teil des Testdatensatzes sind. Dieser Entscheidungstexte dieses Datenausschnittes sind anschließend von zwei Experten der Rechtswissenschaft<sup>43</sup> annotiert worden. Für diese Daten wurden Zitate auf andere Gerichtsentscheidungen, die Nennung der Namen von Richtern und der beschlossene Revisionsausgang (erfolgreich, nicht erfolgreich oder neutral) annotiert. Die Annotation erfolgte mit dem Annotations-Tool *WebAnno*<sup>44</sup>. *WebAnno* ist eine Anwendung, in der Annotatoren mit geringem Aufwand linguistische Daten annotieren können. Dadurch, dass die Anwendung webbasiert ist, muss keine externe Software installiert werden. Es ist außerdem möglich, dass mehrere Personen zeitgleich auf dem selben Datensatz Annotationen durchführen können. *WebAnno* ermöglicht zusätzlich auch das Annotieren von nicht-linguistischen Daten.<sup>45</sup> Um die Daten für den Annotationsprozess in *WebAnno* vorzubereiten, ist es notwendig, die Texte innerhalb einer Entscheidung in eine *tokenisierte* Form zu bringen. Mit tokenisieren ist hier gemeint, dass die Texte in einzelne Segmente auf der Wortebene gebracht werden. Der Satz *Die Kostenentscheidung beruht auf § 97 Abs. 1 ZPO.* würde hier in die folgenden Tokens zerlegt:

Die  
Kostenentscheidung  
beruht  
auf  
§  
97  
Abs  
.   
1  
ZPO  
.

Dieses Format trägt den Namen *CoNLL*-Format.<sup>46</sup> Für *WebAnno* muss anschließend jedem Satz eine fortlaufende Satznummer, und jedem Wort in einem Satz eine fortlaufende Wortnummer zugeordnet werden. Zusätzlich wird zu jedem Wort angegeben, an welcher Stelle im Satz das Wort beginnt und an welcher Stelle es endet (Token-Offsets). Die Tabelle 4.4 demonstriert das abschließende Format, mit dem *WebAnno* die Entscheidungstexte einlesen kann.

<sup>43</sup>Dirk Hartung und Lauritz Gerlach vom Center for Legal Technology and Data Science, Bucerius Law School

<sup>44</sup>Siehe [EdCMMY<sup>+</sup>16]

<sup>45</sup>vgl. [EdCMMY<sup>+</sup>16] S. 76

<sup>46</sup>CoNLL = *Conference on Computational Natural Language Learning*

```

#FORMAT=WebAnno TSV 3.2

#Text=Die Kostenentscheidung beruht auf § 97 Abs. 1 ZPO.
1-1  0-3   Die
1-2  4-22  Kostenentscheidung
1-3  23-29  beruht
1-4  30-33  auf
1-5  34-35  §
1-6  36-38  97
1-7  39-42  Abs
1-8  42-43  .
1-9  44-45  1
1-10 46-49  ZPO
1-11 49-50  .

#Text=Verzugszinsen auf Sozialkassenbeiträge
2-1  51-64  Verzugszinsen
2-2  65-68  auf
...  ...   ...

```

Tabelle 4.4: Tokenisiertes Format für Annotationen in WebAnno

Um die Entscheidungstexte in eine tokenisierte Form zu bringen, wurde das in Abschnitt 4.1.4 vorgestellte Tool OpenNLP verwendet. Innerhalb eines Entscheidungsdokumentes sind allerdings nicht alle Informationen für die Annotation notwendig. Beispielsweise müssen Attribute, wie das Entscheidungsdatum, der Gerichtstyp und das Aktenzeichen, hier nicht tokenisiert werden, da sie für die Annotation keinen Mehrwert bilden. Ferner werden nur der Titel der Entscheidung, der Leitsatz, der Tenor, der Tatbestand, die Entscheidungsgründe, die sonstigen Orientierungssätze und die abweichenden Meinungen für die Annotation berücksichtigt. Für den Annotationsprozess ist zu berücksichtigen, dass die Entscheidungen im Korpus in ihrer Länge stark variieren. Durch die zufällige Auswahl von Entscheidungen aus dem Korpus für den Datenausschnitt können sehr kurze, aber auch sehr lange Entscheidungen ausgewählt werden, wodurch der Aufwand des Annotationsprozesses stark beeinflusst wird. Innerhalb des für diese Arbeit generierten Datenausschnittes enthielt die Entscheidung mit dem längsten Entscheidungstext insgesamt 687 Sätze.<sup>47</sup> Die kürzeste Entscheidung enthielt lediglich einen Satz.<sup>48</sup> Es wird deutlich, dass, um eine hohe Anzahl an annotierten Entscheidungen zu erzielen, entweder mehrere Experten notwendig sind oder genügend Zeit zur Verfügung stehen muss. Stehen, wie in diesem Fall, mehrere Experten zur Verfügung, können außerdem mehre-

<sup>47</sup>vgl. BAG, Urteil vom 05.12.2019 - 2 AZR 240/19, NJW 2020, 1695

<sup>48</sup>vgl. BVerfG, Beschluss vom 01.08.2012 - 1 BvR 2492/08

re Ansätze verfolgt werden: Entweder jede Entscheidung wird von nur einem Experten annotiert, um eine möglichst hohe Anzahl an Entscheidungen annotieren zu können, bzw. möglichst schnell alle Entscheidungen im Trainings- und Testdatensatz zu annotieren. Zudem besteht auch die Möglichkeit, dass jede Entscheidung auch von jedem Experten annotiert wird. Das Annotieren von Entscheidungen durch alle Annotatoren ermöglicht es, unter den Experten ein sog. *inter-annotator-agreement*, also die Übereinstimmung der Experten, zu messen. Bei einer hohen Übereinstimmung kann somit die Glaubwürdigkeit der Grundwahrheit zusätzlich messbar belegt werden.<sup>49</sup> Schlussendlich können, nach Beenden des Annotationsprozesses, die Textdaten exportiert werden, um anschließend die eigene Methodik gegenüber der Grundwahrheit vergleichen zu können.

Um diesen Datensatz in Trainingsdaten und Testdaten aufzuteilen, wurden von den 140 ausgewählten Entscheidungen insgesamt 75% der Entscheidungen für die Trainingsdaten und 25% für die Testdaten verwendet. Der Trainingsdatensatz umfasst somit 105 Entscheidungen, der Testdatensatz umfasst 35 Entscheidungen. Im Rahmen dieser Arbeit wurden insgesamt 140 Entscheidungen annotiert, wobei jede Entscheidung mindestens einmalig und insgesamt 78 Entscheidungen doppelt annotiert wurden. Bei den Entscheidungen, welche von beiden Experten annotiert wurden, konnte dabei nach dem *Krippendorff's alpha* Maß ein *inter-annotator-agreement* von 0.91 erreicht werden. Die Annotationen weisen somit eine hohe Güte auf.

---

<sup>49</sup>vgl. [EdCMMY<sup>+</sup>16] S. 79



## 5 Methodik

Das folgende Kapitel stellt die methodischen Grundlagen der Netzwerkforschung sowie der Sprachtechnologie dieser Untersuchung vor. Es wird hierbei auf die Ziele dieser Arbeit Bezug genommen und anhand dieser mögliche Methoden evaluiert, welche sich für die Ermittlung dieser Ziele eignen. Für die Analyse mit sprachtechnologischen Methoden wird dabei außerdem kritisch betrachtet, ob die vorgestellten Verfahren messbaren Erfolg haben können.

### 5.1 Extraktion der Zitate

Wie in Abschnitt 4.1.3 bereits näher erläutert wurde, werden Gerichtsentscheidungen im Regelfall über ihre Fundstelle zitiert. Allerdings wurde auch festgestellt, dass es Gerichtsbarkeiten gibt, welche entweder neben der Fundstelle zusätzlich das Aktenzeichen zitieren oder nur mit einem Aktenzeichen und ohne Fundstelle auf ein Urteil verweisen. Dies kann beispielsweise auftreten, wenn eine Entscheidung weder in einer amtlichen Sammlung noch in einer Zeitschrift veröffentlicht wird. Um möglichst alle erwähnten Entscheidungen in einem Entscheidungstext finden zu können, müssen für die Extraktion der Zitate, also die Form eines Aktenzeichens, sowie die Form einer Fundstelle gefunden und extrahiert werden. Dies ist grundsätzlich keine triviale Aufgabe. Für die Extraktion der Aktenzeichen besteht hier der Vorteil, dass diese einen schematischen Aufbau verfolgen. Ein Aktenzeichen besteht im wesentlichen aus der Bezeichnung des Spruchkörpers bzw. Senats, einem Registerzeichen, einer Registrierungsnummer sowie einer Jahreszahl. Der Senat gibt an, welche organisatorische Einheit des Gerichtes zuständig ist, das Registerzeichen verweist auf ein bestimmtes Aktenregister des zuständigen Gerichtes, während die Registrierungsnummer eine fortlaufende Nummer innerhalb eines Jahres darstellt. Die Jahreszahl eines Gerichtes beschreibt dazu, in welchem Jahr das Verfahren beim Gericht eingegangen ist. Besitzt eine Geschäftsstelle mehrere Abteilungen, wird dem Aktenzeichen zusätzlich die Abteilungsnummer als arabische Zahl vorangestellt. Liegen Mehrfachakten vor, wird dem Aktenzeichen die Aktennummer als römische Zahl vorangestellt. Es sollte allerdings auch berücksichtigt werden, dass es Aktenzeichen gibt, die von dieser Form abweichen können. Die Aktenzeichen *1 BvR 1024/12* und *VII ZB 9/13* sind hier als Beispiele der gängigen Form zu nennen. Aus dem Aktenzeichen *1 BvR 1024/12* ergibt sich, welcher Senat (1) in einem Verfahren

über Verfassungsbeschwerden (BvR) in der Reihenfolge der fortlaufenden Registrierungsnummer (1024) zuständig ist und dass dieses Verfahren in einem bestimmten Jahr (2012) beim Gericht eingegangen ist. Aus dem Aktenzeichen *VII ZB 9/13* ergibt sich, dass es sich um ein Verfahren über Beschwerden in Zivilsachen und Rechtsbeschwerden handelt (ZB), welches mit fortlaufender Registrierungsnummer (9) im Jahr 2013 eingegangen ist und in die sachliche Zuständigkeit des 7. Senats (*VII*) also Werkvertragsrecht, Architektenrecht und Zwangsvollstreckungsrecht fällt. Die Beispiele zeigen, dass die Struktur eines Aktenzeichens genutzt werden kann, um diese aus dem Text zu extrahieren. Die Fundstellen weisen dabei ebenfalls eine Struktur auf, auf welche in Abschnitt 4.1.3 bereits näher eingegangen wurde. Hier bietet sich also ein regelbasiertes Verfahren an.

Um Fundstellen und Aktenzeichen nun extrahieren zu können, werden *reguläre Ausdrücke* (engl. *regular expressions*) verwendet. Reguläre Ausdrücke ermöglichen es, eine Menge von Zeichenketten anhand von abstrakten Mustern zu definieren. Hierfür können z.B. Platzhalter, konkrete Zeichen und Zeichenmengen, also z.B. alle Ziffern oder alle Großbuchstaben verwendet werden, um ein solches Muster zu bilden. Mit einem generierten Muster besteht im Anschluss die Möglichkeit, Texte nach diesem Muster zu durchsuchen. Wird im Text eine Zeichenkette gefunden, welche dem Muster entspricht, kann dieses extrahiert werden. Dieser Vorgang nennt sich *Pattern Matching*. Ein regulärer Ausdruck, welcher alle Aktenzeichen finden kann, muss somit dem Aufbau des Aktenzeichens nachempfunden werden und alle möglichen Registerzeichen enthalten. Für einen regulären Ausdruck, welcher alle möglichen Fundstellen erkennt, müssten hier ebenfalls alle amtlichen Sammlungen und alle juristischen Fachzeitschriften in den Ausdruck aufgenommen werden. Im Rahmen dieser Arbeit wurde ein regulärer Ausdruck erstellt, welcher einen Großteil der gängigen Strukturen von Aktenzeichen und alle bedeutenden Entscheidungssammlungen sowie die wichtigsten juristischen Fachzeitschriften für Fundstellen abdeckt.<sup>1</sup> Da hier allerdings nur ein Großteil abgedeckt wird, ist damit zu rechnen, dass einige Formen von Aktenzeichen und Fundstellen möglicherweise nicht korrekt oder gar nicht erkannt werden.

Damit sichergestellt werden kann, ob das regelbasierte Verfahren als Methodik Erfolg haben kann, wird es mit den annotierten Daten des in Abschnitt 4.2 vorgestellten Datenausschnittes verglichen. Um die Güte dieses Ansatzes zusätzlich zu stützen, wurde das Verfahren dieser Arbeit außerdem anhand eines unterschiedlichen Testdatensatzes mit den Ergebnissen einer thematisch verwandten Dissertation verglichen. Für diesen Vergleich wird die Veröffentlichung [Cou19] von Corinna Coupette herangezogen. Diese untersucht ein geschlossenes Entscheidungsnetzwerk der Gerichtsentscheidungen des Bundesverfassungsgerichtes seit dessen Gründung auf Basis von Zitaten innerhalb des

---

<sup>1</sup>siehe GitHub: <https://t1p.de/3le5>

Entscheidungstextes. Da Coupette lediglich Entscheidungen des Verfassungsgerichtes untersucht, kann der eben bereits genannte Datenausschnitt nicht für einen Vergleich verwendet werden. Für diese Evaluation wurde deshalb ein weiterer Ausschnitt generiert, welcher 50 zufällig ausgewählte Entscheidungen des Bundesverfassungsgerichtes umfasst. Hier bei wurde lediglich beachtet, dass die Entscheidungen erst nach dem Jahr 2010 veröffentlicht wurden, damit die Entscheidungen in beiden Datensätzen vorhanden sind. Die Tabelle 5.1 beschreibt die Auswertung der Vergleiche.

Verfahren	Precision	Recall	$F_1$
Annotierter Datenausschnitt	0.26	0.66	0.37
Coupette Testdaten	0.97	0.53	0.68

Tabelle 5.1: Evaluation der Extraktion der Zitate

Mit dem regelbasierten Ansatz der Zitatextraktion konnte im Vergleich mit Ground Truth des annotierten nur ein F-Score von 0.37 und im Vergleich zu der Dissertation von Corinna Coupette ein F-Score von 0.68 erreicht werden. Um einordnen zu können, wie diese Werte zustande kommen und warum sie nicht noch höher ausfallen, wurde zudem eine Fehleranalyse durchgeführt. Für den annotierten Datenausschnitt zeigte die Analyse zusammengefasst zwei Aspekte: Zu Beginn zeigt sich, dass der reguläre Ausdruck in einigen Entscheidungen mehr Zitate in den Texten findet als von den Experten annotiert wurden. Eine Analyse mit den Experten ergab, dass dies an vielen Stellen tatsächlich korrekt ist, da in einigen Entscheidungen die Zitate per Fundstelle, sowie über das jeweilige Aktenzeichen erfolgen. Da die Annotatoren in einigen Entscheidungen nur eine der beiden Varianten annotiert haben, sollte dies beim Vergleich berücksichtigt werden. Eine weitergehende Analyse zeigt zudem auch, dass es Zeitschriften gibt, welche im regulären Ausdruck noch nicht enthalten sind und somit nicht extrahiert werden konnten. Zudem bestätigte die Fehleranalyse die Vermutung, dass es Aktenzeichen und Fundstellen mit besonderen Strukturen gibt, welche von dem regulären Ausdruck nicht vollständig erkannt werden. Die eben genannten Aspekte, wurden schließlich für eine

Verfahren	Precision	Recall	$F_1$
Annotierter Datenausschnitt	0.87	0.84	0.85
Coupette Testdaten	0.83	0.64	0.72

Tabelle 5.2: Evaluation nach der Anpassung des Ausdrucks

zweiten Evaluation im regulären Ausdruck angepasst und verbessert. Tabelle 5.2 zeigt nun eine erneute Evaluation mit dem angepassten Ausdruck. Hier konnten deutlich

bessere Werte erzielt werden, wie die  $F_1$ -Scores von 0.85 und 0.72 zeigen. Es zeigt sich also, dass ein regelbasierter Ansatz für die Extraktion der Zitate durchaus gut funktioniert. Bei der Analyse des Vergleiches mit den Daten der Dissertation von Corinna Coupette ist zu erwähnen, dass Coupette lediglich die amtliche Sammlung des Bundesverfassungsgerichtes (BVerfGE) untersucht. Die Vermutung liegt nahe, dass sich auch hier Unterschiede im Extraktionsverfahren ergeben, wodurch Corinna Coupette z.B. teilweise Fundstellen in anderen Formaten extrahiert. Durch diesen Unterschied werden einige Zitate nicht richtig erkannt.

Es bietet sich durchaus an, den regulären Ausdruck in zukünftigen Forschungen weiter zu verfeinern und zu präzisieren, um noch weitere Aktenzeichen und weitere Fundstellen erkennen zu können.

## 5.2 Methoden der Netzwerkforschung

Um anhand eines Netzwerkes nun juristische Netzwerkforschung betreiben zu können, ist es entscheidend einzugrenzen, welche Ebenen eines Netzwerkes genau untersucht werden sollen. In der Netzwerkforschung lassen sich hier drei Untersuchungsebenen beschreiben: (1) die *Mikroebene*, (2) die *Mesoebene* und (3) die *Makroebene*.<sup>2</sup> Auf jeder Ebene können dabei unterschiedliche Fragestellungen betrachtet werden, wodurch auch unterschiedliche Werkzeuge für die Analyse in Frage kommen. Die Mikroebene eines Netzwerkes betrachtet die *lokalen* Strukturen eines Netzwerkes. Es werden hier im genaueren die Eigenschaften einzelner Objekte im Netzwerk, sowie dessen Beziehungen genauer beobachtet. Eine charakteristische Fragestellung ist hier beispielsweise: Welche Elemente im Netzwerk sind besonders wichtig? Außerdem lohnt sich die Betrachtung der Mesoebene. Die Mesoebene untersucht *regionale* Strukturen innerhalb des Netzwerkes. Hier können also Gruppierungen von Objekten und dessen Eigenschaften beleuchtet werden. Als zentrale Fragestellung ist hier anzuführen: Welche Elemente im Netzwerk gehören zusammen? Zuletzt widmet sich die Betrachtung der Makroebene den *globalen* Strukturen eines Netzwerkes. Hier werden Eigenschaften des gesamten Netzwerkes begutachtet. Die zentrale Überlegung ist hier: An welchen Faktoren lässt sich das ganzheitliche Netzwerk besonders beschreiben? Ein Vergleich mit anderen Netzwerken kann sich hier lohnen, um Charakteristiken und Unterschiede des Netzwerkes deutlich zu machen.<sup>3</sup> Auf die Makroebene wird im Rahmen dieser Auswertung jedoch nicht genauer eingegangen, weswegen sie im Folgenden nicht weiter beschrieben wird.

Stattdessen konzentriert sich diese Arbeit auf die Mikroebene und die Mesoebene. Die

---

<sup>2</sup>vgl. [Cou19] S. 71

<sup>3</sup>vgl. [Cou19] S. 72f.

Untersuchungsebenen lassen sich nun auf das Experiment dieser Arbeit beziehen.

### 5.2.1 Mikroebene

Da die Mikroebene sich mit den wichtigsten Objekten eines Netzwerkes befasst, liegt es durch die Netzwerkdefinition dieser Arbeit nahe, die wichtigsten Entscheidungen des Netzwerkes in Betrachtung aller Gerichtsbarkeiten zu bestimmen. Hier ist allerdings eine Definition des Begriffes *Wichtigkeit* von Nöten. Es liegt nahe, die Wichtigkeit einer Entscheidung mit der Häufigkeit der Zitierung gleichzusetzen. Je öfter eine Entscheidung zitiert wird, desto bedeutender wird diese. Eine solche Perspektive ist zunächst schlüssig, allerdings sollte dabei berücksichtigt werden, dass die alleinige Häufigkeit der eingehenden Zitate einer Entscheidung keinen Aufschluss über den eigentlichen Inhalt dieser bietet. Es liegt zwar nahe, dass Entscheidungen eher die Entscheidungen zitieren, welche thematisch mit ihnen verwandt sind, dies ist jedoch ohne die Betrachtung des Entscheidungstextes nicht vollständig belegbar. Grundsätzlich soll die Untersuchung dieser Arbeit herausfinden, wie gut die betrachtete Netzwerkstruktur der Entscheidungen der obersten Gerichtsbarkeiten die eigentliche Rechtsprechung abbildet. In diesem Fall lässt sich die Häufigkeit der Zitate durchaus als geeignetes Mittel ansehen, um wichtige Entscheidungen zu bestimmen, da somit höchstwahrscheinlich Präzedenzfälle und grundlegende Urteile zu verschiedenen Rechtsthematiken häufig zitiert werden, welche sich für viele spätere Urteile zu einem Maßstab entwickelt haben und somit von größerer Bedeutung sind.

Um die Wichtigkeit eines Objektes im Netzwerk zu bestimmen, gibt es nun verschiedene Methoden, welche sich für eine Analyse eignen. Als erste und einfachste Möglichkeit ist hier die Gradzentralität (engl. *degree centrality*) zu nennen. Sie entspricht im Wesentlichen der Anzahl der Beziehungen, an denen ein Objekt direkt beteiligt ist. Da diese Arbeit einen gerichteten Graphen betrachtet, wird hier zusätzlich der *Eingangsgrad* (engl. *indegree*) und der *Ausgangsgrad* (engl. *outdegree*) unterschieden. Der Eingangsgrad eines Objektes ist definiert als die Anzahl aller auf das Objekt zeigenden Zitate, der Ausgangsgrad als die Anzahl aller vom Objekt ausgehenden Zitate. Der Eingangsgrad misst somit für jede Entscheidung, wie oft sie zitiert wurde, wobei der Ausgangsgrad misst, wie oft eine Entscheidung selbst andere Entscheidungen zitiert. Gradzentralitäten verfügen jedoch allerdings über den Nachteil, dass allen Entscheidungen die gleiche Bedeutung zugemessen wird. Sinnvoller erscheint es hier, die Gewichtung eines Objektes im Netzwerk zusätzlich davon beeinflussen zu lassen, wie wichtig ihre zitierenden Nachbarobjekte sind. Es wird hier von *Rückkopplungszentralitäten* (engl. *feedback centralities*) gesprochen.<sup>4</sup>

---

<sup>4</sup>vgl. [Cou19] S. 135

Ein Verfahren, welches diesen Aspekt berücksichtigt, ist der in Abschnitt 2.1.1 bereits vorgestellte PageRank Algorithmus. Verschiedene themenverwandte Veröffentlichungen bestätigen bereits, dass PageRank auch im Kontext von juristischer Netzwerkforschung erfolgreich eingesetzt werden kann, um innerhalb eines Korpus die wichtigsten Entscheidungen zu bestimmen.<sup>5</sup> Um die Güte sicherzustellen, wird das Verfahren in dieser Arbeit dennoch anhand einer qualitativen Analyse überprüft. Innerhalb dieser Analyse werden einem rechtswissenschaftlichen Experten die zehn Entscheidungen mit dem höchsten PageRank des Korpus vorgelegt. Der Experte überprüft anschließend, ob es sich hier um zentrale Entscheidungen oder Präzedenzfälle handelt.

Um eine möglichst umfassendes Bild der Wichtigkeit von Objekten im Netzwerk erhalten zu können, sollte zuletzt noch ein weiteres Verfahren betrachtet werden. Denn Gradzentralitäten und Rückkopplungszentralitäten befassen sich lediglich mit lokalen Informationen. Dies bedeutet, dass entweder lediglich Objekte selbst betrachtet werden (Gradzentralität) oder Objekte und Beziehungen der unmittelbaren Nachbarschaft eines Objektes Einfluss auf dessen Gewichtung nehmen können (Rückkopplungszentralität). Es werden also globale Faktoren des Netzwerkes bei der Bestimmung der Wichtigkeit eines Objektes vernachlässigt. Zentralitätsmaße, welche sich diesen Problemen annehmen, werden als *Positionszentralitäten* bezeichnet. Hier gibt es zwei mögliche Verfahren, welche globale Faktoren berücksichtigen: Die *Nähezentralität* (engl. *closeness centrality*) und die *Zwischenzentralität* (engl. *betweenness centrality*).<sup>6</sup> Die Nähezentralität misst hierbei die Nähe eines Objektes zu allen anderen Objekten im Netzwerk, wobei hier mit *Nähe* der Wegabstand innerhalb des Netzwerkes gemeint ist. Dieser definiert sich durch die Anzahl an Kanten, welche von einem Objekt durchlaufen werden müssen, um zu dem untersuchten Objekt zu gelangen.<sup>7</sup> Die Zwischenzentralität befasst sich dagegen damit, inwieweit ein Objekt im Netzwerk zwischen den anderen Objekten im Netzwerk steht. Somit können im Prinzip Vermittlerobjekte gefunden werden, über die andere Objekte im Netzwerk verbunden werden.

Neben dem Herausarbeiten der Wichtigkeit einer Entscheidung werden im Rahmen dieser Arbeit außerdem die regionalen Strukturen des Netzwerkes untersucht. Hierfür soll nun ermittelt werden, welche Entscheidungen auf Basis der Beziehungsstruktur zusammengehören und gruppiert werden können.

---

<sup>5</sup>siehe [DL14] S. 677 und [WRK11] S. 110

<sup>6</sup>vgl. [Cou19] S. 139f.

<sup>7</sup>Hier wird im Allgemeinen vom kürzesten möglichen Weg von einem Objekt zum untersuchten Objekt ausgegangen. Siehe beispielsweise: [NS00].

### 5.2.2 Mesoebene

Analog zu der Definition der Wichtigkeit auf der Mikroebene ist es für die Mesoebene von grundsätzlicher Bedeutung, wie der Begriff *Zusammengehörigkeit* definiert wird, um überhaupt Gruppierungen vornehmen zu können. Hierfür ist das Ziel der Gruppierung bedeutend. Innerhalb dieser Arbeit wird eine *Fokussierung* beabsichtigt, bedeutet es werden kleinere Subnetzwerke erzeugt, welche anschließend genauer betrachtet werden können. Ziel dieser Untersuchung ist im Idealfall das Gruppieren von Gerichtsentscheidungen, welche sich häufig untereinander zitieren und somit höchstwahrscheinlich einen inhaltlichen Zusammenhang besitzen. Anhand der Untersuchung soll auch geprüft werden, ob sich Entscheidungen eines Gerichtes eher untereinander zitieren oder ob auch häufige gerichtsübergreifende Zitate stattfinden. Ersteres würde beispielsweise dadurch erkennbar, dass die Entscheidungen der einzelnen Gerichtsbarkeiten auch einzelne Gruppierungen bilden. Um innerhalb des Korpus im Rahmen einer Fokussierung verschiedene Gruppierungen zu bestimmen, wird der in Abschnitt 2.1.2 vorgestellte Algorithmus Chinese Whispers verwendet.

Um die Güte des Ansatzes mit dem Chinese Whispers Algorithmus zu evaluieren, wird hier schließlich eine qualitative Analyse der Gruppierungen durch einen Rechtswissenschaftsexperten durchgeführt. Diese qualitative Analyse erfolgt im Rahmen eines *Blindtests*, bedeutet, dem Experten werden lediglich zehn willkürlich gewählte Entscheidungen einer Gruppierung vorgelegt. Es wird dabei keine Aussage darüber getroffen, wie diese Entscheidungen miteinander in Beziehung stehen, ob sie direkt miteinander in Beziehung stehen und wie viele Knoten letztendlich der gesamten Gruppierung angehören. Der Experte evaluiert anschließend, ob diese Entscheidungen einen inhaltlichen Zusammenhang haben und gibt jedem Ausschnitt einer Gruppierung, wenn möglich, ein *Cluster Label*, mit dem dieser Ausschnitt beschrieben werden kann. Dieses Verfahren wird an insgesamt zehn zufällig gewählten Gruppierungen durchgeführt.

Neben einer Fokussierung kann das Ziel einer Untersuchung auch eine *Komprimierung* oder eine *Strukturierung* sein. Während eine Komprimierung zum Ziel hat, die zusammengehörigen Gruppen zu einem Objekt im Netzwerk zusammenzufassen, ermöglicht die Strukturierung, dem Netzwerk Informationen hinzuzufügen, welche dessen Organisation und Aufbau näher beschreiben.<sup>8</sup> Dies ist allerdings nicht Gegenstand dieser Arbeit und wird daher nicht genauer betrachtet.

---

<sup>8</sup>vgl. [Cou19] S. 143

### 5.3 Methoden der Sprachtechnologie

Auch für die sprachtechnologische Analyse ist eine genauere Eingrenzung notwendig, welche Bereiche für die Rechtswissenschaft einen Nutzen haben können und welche Gegenstände dafür genau untersucht werden sollen. Recht, welches zumeist in textlicher Form untersucht werden kann, weist dabei grundsätzlich verschiedene Möglichkeiten auf, um sprachtechnologische Methoden einzusetzen. Hier seien im Folgenden einige Beispiele exemplarisch zu nennen, in denen NLP im Recht einen großen Nutzen hat. Es kann zum Beispiel ein Ansatz verfolgt werden, welcher sich mit der Relevanz von juristischen Dokumenten befasst. So kann versucht werden, die Relevanz einer Entscheidung zu messen und somit anderen Entscheidungen gegenüber zu vergleichen. Eine weitere Möglichkeit besteht darin, auf Basis der Analyse von vergangenen Entscheidungsdokumenten, ein System zu entwickeln, welches das Erstellen von neuen Dokumenten teilweise automatisiert. Denkbar wäre auch, Informationen aus Entscheidungen zu nutzen, um Juristen bei der zukünftigen Entscheidungsfindung zu unterstützen. Diese Ansätze teilen im Wesentlichen einen wichtigen Schwerpunkt, denn für ihre Realisierung ist das automatisierte Finden und Extrahieren von Informationen aus textbasierten Dokumenten notwendig. In der Computerlinguistik wird hier von *Informationsextraktion* (engl. *information extraction*) gesprochen. In dieser Arbeit wird ein Fokus auf die Methoden dieses Fachgebietes gelegt. Es sollen also die juristischen Entscheidungsdokumente untersucht werden, um relevante Informationen aus ihnen zu extrahieren.

Im Hinblick auf den vorliegenden Datensatz liegen bereits einige wichtige Informationen wie z.B. das Aktenzeichen, der Spruchkörper und das Entscheidungsdatum vor. Dazu kann eine Betrachtung von juristischen Entscheidungsträgern durchaus relevant sein. Eine Extraktion der Richter, welche an einer Entscheidung beteiligt sind oder welche innerhalb einer Entscheidung genannt werden, bietet die Möglichkeit, Datensätze neu zu strukturieren oder schneller gezielt durchsuchbar zu machen. Außerdem ermöglicht sie eine Auswertung des individuellen Einflusses einzelner Personen auf Parameter der Entscheidung, wie deren Länge oder deren Bedeutung. Das ist von besonderer Relevanz, da sich die Zusammensetzung der einzelnen Senate über die Zeit ändert und dadurch der Einfluss einzelner Personen verloren geht, sofern man nur die Spruchkörper anhand des Aktenzeichens betrachtet. Wirft man nun einen Blick auf den Bereich der Netzwerkforschung und versucht dieses Fachgebiet mit der Sprachtechnologie zu verbinden, ergeben sich außerdem weitere Möglichkeiten, um ein Netzwerk mit den hier gewonnenen Informationen zu erweitern. Hier könnten exemplarisch Korrelationen von Richtern mit Entscheidungsverweisen untersucht werden. Gibt es Richter, dessen Entscheidungen signifikant häufiger zitiert werden, als die Entscheidungen anderer Richter? Das Extrahieren von Entitäten, also z.B. Personen, Organisationen oder Orten aus einem natürlichsprachlichen Text wird als *Named Entity Recognition* bezeichnet.

### 5.3.1 Extraktion der Richter

Um nun speziell Richter als eine Entität aus den Entscheidungstexten extrahieren zu können bestehen mehrere mögliche Verfahren. Hier tritt grundsätzlich das Problem auf, dass die Namen von Richtern sich in einer Entscheidung grundsätzlich nicht von den Namen anderer genannter Personen ohne juristischen Bezug unterscheiden lassen. Zumeist kann lediglich durch den Kontext, mit Schlüsselworten wie z.B. *Richter* oder *Verfassungsrichter* auf einen Richternamen geschlossen werden. Im Übrigen gibt es im deutschen Raum häufig Nachnamen, welche mehrmalig auftreten oder welche z.B. gleichbedeutend mit einer Ortschaft sind. Die Anonymisierung der Gerichtsentscheidungen trägt zu einer geringeren Verwechslung von am Verfahren beteiligten Personen, gegenüber Richtern bei, da jegliche personenbezogene Daten aus einer Entscheidung entfernt werden. Dies gilt allerdings nicht bzw. nur eingeschränkt für Personen des öffentlichen Lebens. So wird beispielsweise der Bundesminister des Innern, für Bau und Heimat, Horst Seehofer, als öffentliche Person nicht anonymisiert und kann somit ungewollt als Person erkannt und extrahiert werden.<sup>9</sup> Eine entscheidende Aufgabe ist es also, Richternamen von anderen auftretenden Personen innerhalb des Entscheidungstextes zu unterscheiden.

Zum Vorteil dieser Untersuchung stehen alle Richternamen, welche den obersten Gerichtsbarkeiten der Bundesrepublik Deutschland und dem Bundesverfassungsgericht seit ihrer Gründung angehört haben als öffentlich zugängliche Liste zur Verfügung.<sup>10</sup> Für das Bundespatentgericht ist ebenfalls eine Liste vorhanden.<sup>11</sup> Hier wird allerdings nur eine geringe Anzahl an Richternamen aufgeführt. Das Gericht selbst hält jeweils eine aktuelle Liste vor, historisch vollständige Daten sind nicht verfügbar.<sup>12</sup> Ein grundsätzlich mögliches Verfahren ist hier also die Liste der Richter als *Wörterbuch* (engl. *Dictionary*) zu verwenden, schließlich automatisiert Wort für Wort durch den Entscheidungstext zu iterieren und jedes Wort mit dem Wörterbuch abzugleichen. Der Text wird dafür tokenisiert und in das CoNLL-Format überführt.<sup>13</sup> Taucht ein Richtername im Text auf wird dieser beim Abgleich erkannt und das Wort wird im Text mit der Klasse Person annotiert. Dies geschieht nach einem *Begin, Inside, Outside*-Schema (BIO-Schema), in welchem der Beginn einer Entität als Beginn der Person, also *B-Per* annotiert wird.<sup>14</sup> Sollte der Name mehrere Wörter umfassen, z.B. durch genannte Vornamen, Zweitnamen und Nachnamen, werden die folgenden Wörter als *Inside* der Person, also *I-Per* annotiert. Wörter, welche keiner Entität angehören, erhalten ein *O* und gehören damit

<sup>9</sup>vgl. BVerfG, Urteil vom 09.06.2020 - 2 BvE 1/19, NJW 2020, 2096

<sup>10</sup>siehe [https://de.wikipedia.org/wiki/Liste\\_der\\_deutschen\\_Bundesrichter](https://de.wikipedia.org/wiki/Liste_der_deutschen_Bundesrichter)

<sup>11</sup>siehe [https://de.wikipedia.org/wiki/Kategorie:Richter\\_\(Bundespatentgericht\)](https://de.wikipedia.org/wiki/Kategorie:Richter_(Bundespatentgericht))

<sup>12</sup>siehe <https://www.bundespatentgericht.de/SharedDocs/Downloads/DE/dasGericht/dienstaltersliste.pdf>

<sup>13</sup>siehe Abschnitt 4.2

<sup>14</sup>vgl. [RR09] S. 150

der Klasse *Outside* an. Bei diesem Verfahren sollte allerdings berücksichtigt werden, dass es Richternamen geben kann, welche beispielsweise als Orte im Text vorkommen können.

Als eine weitere Möglichkeit können schließlich Algorithmen für diese Aufgabe verwendet werden, welche speziell für Named Entity Recognition Tasks entwickelt wurden. Im Rahmen dieser Arbeit wird für die Aufgabe der Richterextraktion der *GermaNER* als NER Modell verwendet.<sup>15</sup> Der *GermaNER* ist ein vortrainiertes NER Modell für deutsche Sprache, welches mit 31.300 Sätzen des GermEval 2014 Datensatzes trainiert wurde. Der *GermaNER* annotiert dabei Personen, Orte, Organisationen und andere Entitäten (*Other*). Der *GermaNER* benötigt als Eingabe ein Format, welches dem CoNLL Format ähnelt. Das Eingabedokument enthält pro Zeile jeweils ein Token. Die Sätze werden dabei durch eine leere Zeile getrennt. Die Annotation erfolgt dabei dem BIO-Schema, wobei Personen als *PER*, Organisationen als *ORG*, Orte als *LOC* und andere Entitäten als *OTH* annotiert werden. Mit der Entität *Other* ist hier gemeint, dass es Entitäten geben kann, welche zwar als Entität erkannt werden, jedoch keiner der vorhandenen Klassen eindeutig zugeordnet werden können.

Um die Güte dieser verschiedenen Verfahren zu evaluieren, werden sie auf die in Abschnitt 4.2 bereits vorstellten rohen Daten ausgeführt und anschließend mit den annotierten Daten, also der Ground Truth, verglichen. Das Verfahren, welches die Richterliste als Wörterbuch nutzt, verwendet keine maschinellen Lernalgorithmen und bedarf daher keines weiteren Trainings. Der *GermaNER* kann aufgrund seines Vortrainings bereits ohne Modifikation für NER Tasks verwendet werden. Es bestünde auch die Möglichkeit den *GermaNER* mit den annotierten Trainingsdaten zu trainieren, dies wird jedoch im Rahmen dieser Arbeit nicht weiter betrachtet.

Die Tabelle 5.3 zeigt die Evaluation der beiden Verfahren in einer Gegenüberstellung. Bei der Betrachtung der Werte fällt auf, dass der  $F_1$ -Score des *GermaNER* sehr nied-

Verfahren	Precision	Recall	$F_1$
Wörterbuch	0.62	0.65	0.64
GermaNER	0.44	0.57	0.49

Tabelle 5.3: Evaluation der Verfahren zur Richterextraktion

rig ausfällt. Eine genauere Untersuchung der Auswertungen zeigt, dass sich hier die Vermutung bestätigt, dass Richternamen teilweise als Ortschaft erkannt werden und somit der Klasse *LOC* zugeordnet werden, da der Richternamen dem Namen eines Ortes gleicht. Genauso gibt es, wie im Wörterbuchverfahren bemerkt, z.B. Richter mit

<sup>15</sup>vgl. [BMPB15]

den Nachnamen *Winter* oder *Wand*. Durch die Mehrdeutigkeit dieser Wörter erklärt es sich, dass es schwierig ist, mit dieser Methode eine möglichst hohe Anzahl an korrekten Extraktionsergebnissen zu erzielen, da hier häufig zu viele Wörter extrahiert werden. Die Methode bestätigt umso mehr, dass es in Entscheidungstexten notwendig ist, den Kontext zu berücksichtigen um Richter präzise extrahieren zu können. Bei der Extraktion der Entitäten mit einem Wörterbuch bestätigt sich zudem, dass die herangezogenen Richterlisten nicht vollständig sind, sodass einige Richter nicht erkannt werden.

Bei der Extraktion der Richter bestünde auch die Möglichkeit, die Extraktion mit einem Transformer Modell durchzuführen, um zu versuchen mit modernen Technologien noch präzisere Messergebnisse zu erwirken. Dies wird jedoch im Hinblick auf den Umfang dieser Arbeit nicht genauer betrachtet, soll jedoch Teil von zukünftigen Forschungen werden.

#### 5.3.2 Klassifikation des Revisionsausgangs

Neben einer Extraktion von Informationen aus juristischen Texten kann auch eine semantische Analyse neue Erkenntnisse ermöglichen. Interessant ist hier eine Betrachtung des Revisionsausgangs einer Entscheidung.

Die Justiz dient der Schaffung von Rechtsfrieden durch möglichst nachvollziehbare, gerechte Entscheidungen. Die Justizgrundrechte gewährleisten, dass sich die Entscheidungen nur am Gesetz orientieren. Aber auch im gerichtlichen Entscheidungsprozess können Fehler passieren oder unterschiedliche Auffassungen bestehen. Um die Beteiligten vor möglichen Fehlern zu schützen kann daher gegen eine Entscheidung eine *Revision* eingelegt werden. Die Revision stellt ein Rechtsmittel dar, mit dem eine Entscheidung an eine weitere Instanz weitergegeben wird, um diese Entscheidung auf Rechtsfehler zu überprüfen. Anders als bei der *Berufung*, werden bei der Revision allerdings nicht mehr die Umstände des Falles untersucht, sondern lediglich überprüft ob bei der Entscheidung der Vorinstanz rechtliche Fehler aufgetreten sind. Eine Revision wird von der folgenden Instanz als *positiv* oder *erfolgreich* eingestuft, wenn die nachfolgende Instanz tatsächlich einen Rechtsfehler der Vorinstanz erkennt. Wird eine Revision als *negativ* oder *nicht erfolgreich* eingestuft, bedeutet dies, dass keine Rechtsfehler festgestellt werden konnten. Da im Korpus auch andere Verfahrensarten (wie beispielsweise Kostenstreitigkeiten) enthalten sind, die sich nicht mit Rechtsfehlern beschäftigen, können Entscheidungen auch als *neutral* eingestuft werden. In einigen Kontexten kann die Revision auch als *teilweise erfolgreich* oder *teilweise nicht erfolgreich* eingestuft werden. Dies wird in dieser Arbeit allerdings nicht genauer untersucht.

Andere Verfahrensarten werden im juristischen Sprachgebrauch nicht immer als Revisionen bezeichnet, die Terminologie wird in dieser Untersuchung aber einheitlich verwendet. Für die quantitative Untersuchung sind die unterschiedlichen Bezeichnungen

nicht ergebnisrelevant.

Innerhalb dieser Arbeit wird ein Transformer Modell verwendet, um anhand des Entscheidungstextes klassifizieren zu können, ob die Revision der Entscheidung als positiv, negativ oder neutral eingestuft wurde. Für diese Aufgabe wird der Transformer *ELECTRA* verwendet. Das Modell soll dabei einen Satz als Eingabesequenz erhalten und anschließend ausgeben, ob dieser Satz im Bezug auf die Revision als positiv, negativ, neutral oder irrelevant zu bewerten ist. Da die Gerichtsentscheidungen in deutscher Sprache vorliegen, muss hier ein Transformer Modell verwendet werden, welches deutsche Sätze verarbeiten kann. Für diese Arbeit wurde das *German Electra Uncased* Model verwendet.<sup>16</sup> Dieses Modell wurde im August 2020 veröffentlicht und ist, nach Angabe der Autoren, zum Zeitpunkt der Veröffentlichung das beste deutsche NLP Modell, gemessen an der Performance verschiedener Evaluationsmetriken.<sup>17</sup> Der Transformer wurde unter anderem auf Common Crawl Daten, Wikipedia Artikel und Untertitel vortrainiert. Für den Einsatz in dieser Arbeit wird das Transformer-Modell dann anhand der in Abschnitt 4.2 vorgestellten annotierten Trainingsdaten trainiert. Zusätzlich zu diesem Datensatz werden noch etwa 2.200 Sätze aus einer Vorarbeit<sup>18</sup> verwendet, um die Präzision des Modells zusätzlich zu verbessern. Die Abbildung 5.1 veranschaulicht

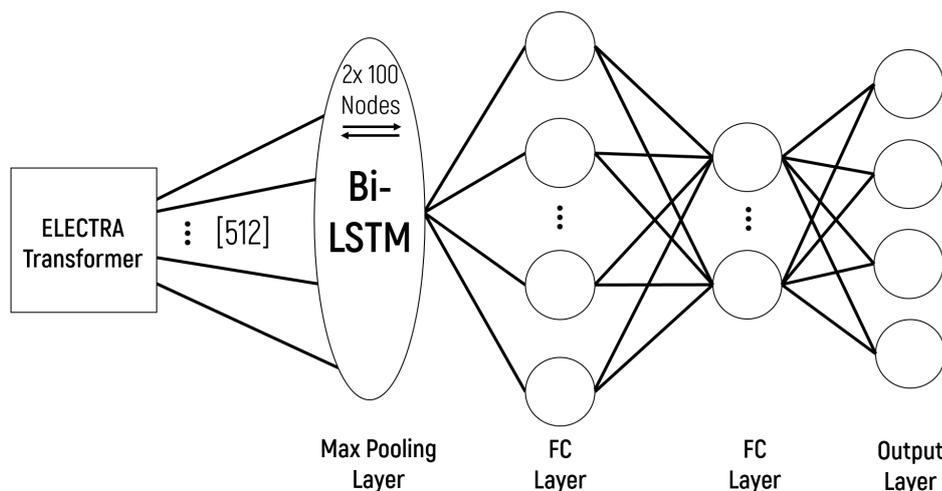


Abbildung 5.1: Transformer Architektur

die Architektur des in dieser Arbeit verwendeten Modells. Zu Beginn wird ELECTRA die Inputsequenz übergeben. ELECTRA encodiert anschließend diese Sequenz und gibt diese mit 512 Nodes anschließend an eine Bidirektionale LSTM-Schicht, welche jeweils

<sup>16</sup>vgl. BEYOND BERT – Challenges and Potentials in the Training of German Language Models, <https://youtu.be/cxgrTd2AQis> (Die wissenschaftliche Veröffentlichung ist noch im Review).

<sup>17</sup>vgl. <https://huggingface.co/german-nlp-group/electra-base-german-uncased>

<sup>18</sup>siehe [RHS<sup>+</sup> 18]

100 Nodes in jede Richtung aufweist. Die LSTM-Schicht wird anschließend trainiert und an die Klassifikation angepasst. Das LSTM berechnet schließlich eine hochdimensionale Repräsentation der Sequenz, welche anschließend in einer *Max Pooling-Schicht* auf eine Dimension aggregiert wird, damit diese für die kommenden Schichten verarbeitbar sind. Schließlich wird die Sequenz an ein großes *fully connected layer* mit 100 Nodes und schließlich in ein kleineres *fully connected layer* mit 50 Nodes übergeben. Diese beiden Schichten verwenden dabei *Rectifier* als ihre Aktivierungsfunktion. Die Aktivierungsfunktion von *Rectified linear units* (ReLU) ist definiert als  $f(x) = \max(0, x)$ .<sup>19</sup> Zuletzt wird die Eingabesequenz an das Output Layer übergeben. Hierfür wird eine *Softmax Funktion* als Aktivierungsfunktion angewandt. Diese sorgt dafür, dass eine Wahrscheinlichkeitsverteilung über die vier Kategorien ausgegeben werden kann.<sup>20</sup> Das Output Layer besteht also schließlich aus vier Nodes, welche jeweils die Klassen positiv, neutral, negativ und irrelevant abbilden. Dort wird schließlich ermittelt, welche Kategorie die höchste Aktivierung aufweist. Für das Modell wird zudem der Optimizer *Adam* verwendet.<sup>21</sup>

Anhand der in diesem Kapitel beschriebenen Verfahren kann nun eine netzwerktechnische und eine sprachtechnologische Auswertung auf den Datensatz erfolgen.

---

<sup>19</sup>vgl. [GBB11] S. 315

<sup>20</sup>vgl. [GP17] S. 1

<sup>21</sup>vgl. [KB17]



## 6 Ergebnisse

Im folgenden Kapitel werden innerhalb eines Experiments die Entscheidungen des Bundesverfassungsgerichtes, der obersten Gerichtshöfe des Bundes, sowie des Bundespatentgerichtes mit den in Kapitel 5 benannten Methoden untersucht. Es sollte allerdings berücksichtigt werden, dass die durchgeführten Analysen nur einen Teil der insgesamt betrachteten Rechtsprechung darstellen können. Da im Rahmen der Datenerhebung und der Datenaufbereitung nur ein Teil der Mapping Daten der Fundstellen zu einem Aktenzeichen vorliegen, um eine Fundstelle in ein Aktenzeichen zu überführen, kann nicht allen extrahierten Zitaten ein Ziel zugeordnet werden, wodurch einige Beziehungen im Netzwerk nicht dargestellt werden können. Hinzu kommt, dass im Rahmen der Netzwerkdefinition einige Entscheidungen getroffen wurden, welche die Analyse beeinflussen. So kann beispielsweise eine Analyse eines geschlossenen Netzwerkes andere Ergebnisse aufweisen, als diese Analyse, welche ein offenes Netzwerk betrachtet. Die Ergebnisse dieser Arbeit sollten mit diesem Bewusstsein interpretiert werden.

Innerhalb des Entscheidungsnetzwerkes sind die Mikroebene und die Mesoebene untersucht worden. Innerhalb der Mikroebene werden dabei lokale Strukturen betrachtet, um besonders zentrale Entscheidungen innerhalb des Netzwerkes zu ermitteln. Anhand der Gradzentralität lässt sich dabei eine erste Auswertung vornehmen. Diese wurde einem Rechtswissenschaftsexperten zur qualitativen Analyse vorgelegt. Die Tabelle 6.1 stellt die zehn Entscheidungen mit dem größten Eingangsgrad dar. Hier werden also nun Entscheidungen aufgelistet, welche besonders viele eingehende Kanten besitzen. Es ist zu erkennen, dass die Entscheidung mit dem höchsten Eingangsgrad insgesamt 349 Eingangskanten aufweist, somit also 349-mal auf sie verwiesen wurde. Es handelt sich hier um die Entscheidung *1 BvR 986/91 - Verletzung des Anspruchs auf Rechtliches Gehör nach Art. 103 Abs. 1 GG, wenn das Gericht auf einen wesentlichen Teil des Tatsachenvortrags nicht eingeht*. Dass diese Entscheidung regelmäßig zitiert wird ist hier durchaus plausibel, wenn in einem Verfahren ein Gericht die Ausführungen einer Partei nicht ausreichend berücksichtigt hat. Auch die Entscheidung *1 BvR 1383/90 - Verletzung des Anspruchs auf Rechtliches Gehör nach Art. 103 Abs. 1 GG, wenn das Gericht Anforderungen an den Tatsachenvortrag einer Partei stellt, mit denen diese nicht rechnen konnte*, erscheint hier als ein plausibles Ergebnis, da auch diese Entscheidung regelmäßig zitiert wird, wenn in einem Urteil, gegen dass sich ein Betroffener mit einer Verfassungsbeschwerde wendet, die Anforderungen an dessen Vortrag im Vorverfahren

## 6 Ergebnisse

Rang	Aktenzeichen	Gerichtstyp	Eingangsgrad
1	1 BvR 986/91	BVerfG	349
2	1 BvR 1383/90	BVerfG	214
3	V ZR 110/60	BGH	160
4	1 BvR 1291/85	BVerfG	153
5	2 BvR 2333/08, 2 BvR 2365/09 2 BvR 571/10	BVerfG	152
6	1 BvL 1/09, 1 BvL 3/09, 1 BvL 4/09	BVerfG	148
7	1 BvR 1621/94	BVerfG	140
8	2 1 BvR 857/07	BVerfG	140
9	2 BvR 2029/01	BVerfG	137
10	1 BvR 1916/09	BVerfG	137

Tabelle 6.1: Die häufigsten zitierten Entscheidungen des Korpus nach Eingangsgrad

überstrapaziert wurden. Bei der Entscheidung auf dem vierten Rang handelt es sich zudem um eine Grundsatzentscheidung von verfassungsgerichtlichen Verfahren. Auf der Grundlage dieser Entscheidung wird die Vergütung von Rechtsanwälten berechnet, sodass dieser Gegenstand regelmäßig in Verfahren beantwortet werden muss. Hier ist es durchaus nachvollziehbar, dass diese Entscheidung häufig zitiert wird.

Bei der genaueren Betrachtung der Auflistung des Eingangsgrades bestätigt sich grundsätzlich in vielen Fällen ein intuitives Judiz. Es ist zu erkennen, dass populäre Verfahren wie beispielsweise zur Sicherungsverwahrung (*2 BvR 2333/08, 2 BvR 2365/09, 2 BvR 571/10*) häufig zitiert werden. Eine genauere Betrachtung der Rangordnung zeigt jedoch, dass diese allerdings nicht überall nachvollziehbar ist. Dass beispielsweise eine Entscheidung zum Hochschulrechtlichen Beamtenrecht (*1 BvR 1621/94*) als zentralere Entscheidung eingestuft wird als eine Entscheidung zur Grundrechtsberechtigung europäischer juristischer Personen (*1 BvR 1916/09*), lässt sich normativ zunächst nicht eindeutig bestimmen. Es ist hier durchaus möglich, dass ein Hinzufügen von Zitaten, welche durch die aktuell vorhandenen Fundstellendaten nicht erfasst werden konnten, dieses Bild verändern könnte. Grundsätzlich kann bei einer Auswertung des Eingangsgrades gesagt werden, dass die Entscheidungen einen sinnvollen Maßstab für zentrale Entscheidungen bilden. Um hier auf einer detaillierteren Ebene die Zentralität zu bestimmen, reicht dieses Verfahren allerdings nicht aus.

Neben dem Innengrad kann auch der Außengrad der Entscheidungen betrachtet wer-

Rang	Aktenzeichen	Gerichtstyp	Ausgangsgrad
1	B 6 KA 25/19 R	BSG	105
2	4 AZR 549/08 (A)	BAG	102
3	10 AZR 121/18	BFH	101
4	4 AZR 549/08	BAG	97
5	XI ZR 355/18	BGH	94
6	10 AZR 330/16	BAG	93
7	2 BvL 1/12	BVerfG	90
8	6 AZR 49/12	BAG	90
9	1 StR 159/17	BGH	90
10	10 AZR 330/16 (A)	BAG	89

Tabelle 6.2: Die am häufigsten zitierenden Entscheidungen des Korpus nach Ausgangsgrad

den. Die Tabelle 6.2 stellt hier die Entscheidungen dar, welche am häufigsten andere Entscheidungen zitiert haben. Werden die hier dargestellten Entscheidungen genauer betrachtet, lässt sich erkennen, dass es sich – im Hinblick auf die Länge der Entscheidungen – um sehr umfangreiche Entscheidungen handelt. Innerhalb des Entscheidungstextes kann dabei eine große Anzahl an Zitaten zu anderen Entscheidungen festgestellt werden. Grundsätzlich ist hier allerdings einzuräumen, dass sich anhand dessen keine besondere Bedeutung ableiten lässt. Entscheidungen, welche selbst viele Entscheidungen zitieren, könnten dabei vielmehr als besonders vielfältige Entscheidungen angesehen werden, welche diverse Themen betreffen. Eine inhaltliche Auswertung der Entscheidungen zeigt jedoch, dass es keine nachvollziehbare Erklärung für den hohen Ausgangsgrad gibt. In diesem Fall sollten zukünftig weitere Metriken evaluiert werden.

Dem Eingangsgrad kann außerdem ein Verfahren mit Rückkopplungszentralitäten gegenüber gestellt werden, um hier in der Bedeutung einzelner Entscheidungen zu unterscheiden. In der Tabelle 6.3 sind die zehn zentralsten Entscheidungen des Netzwerkes nach der Berechnung des PageRank Algorithmus' aufgelistet. Der PageRank bildet grundsätzlich ebenfalls die Bedeutung einer Entscheidung anhand ihres Eingangsgrades ab. Dies ist bei dem Vergleich mit der Tabelle des Eingangsgrades sofort zu erkennen, da hier die ersten beiden Entscheidungen identisch sind. Innerhalb der zehn Entscheidungen mit dem höchsten PageRank finden sich sogar insgesamt vier Entscheidungen, welche auch in der Rangliste des Eingangsgrades vorkommen. Bei weiterer Betrachtung der beiden Tabellen fällt zudem auf, dass ein Großteil der Entscheidungen in den Ranglisten dem Bundesverfassungsgericht angehören. Dies ist besonders im Hinblick auf den Anteil der Entscheidungen des BVerfG am Korpus interessant, da das BVerfG wie in Abschnitt 4.1.5 dargestellt den geringsten Anteil am Korpus hat. Da diese Arbeit ein offenes Netzwerk untersucht, war im Vorfeld der Analyse bereits davon auszugehen, dass ein Teil dieser Entscheidungen nicht Teil des Korpus sein wird. Eine Be-

## 6 Ergebnisse

Rang	Aktenzeichen	Gerichtstyp	PageRank
1	1 BvR 986/91	BVerfG	0.0005620087
2	1 BvR 1383/90	BVerfG	0.0003250027
3	GrS 4/82	BFH	0.0003101583
4	1 BvR 857/07	BVerfG	0.00028422437
5	1 BvR 1621/94	BVerfG	0.00023486663
6	2 BvR 527/99, 2 BvR 1337/00, 2 BvR 1777/00	BVerfG	0.00020406026
7	V ZB 172/09	BGH	0.00020373722
8	2 BvR 167/02	BVerfG	0.00019761976
9	3 GrS 2/92	BFH	0.00019530237
10	1 BvR 1934/93	BVerfG	0.00019035217

Tabelle 6.3: Die zehn zentralsten Entscheidungen des Korpus nach PageRank

trachtung der Entscheidungen der Rangliste bestätigt diese Vermutung. Dies liegt mit daran, dass der Datensatz nur Entscheidungen seit dem Jahr 2010 enthält und ältere Entscheidungen häufiger zitiert werden können. Der hohe PageRank dieser Entscheidungen gibt allerdings auch einen entscheidenden Hinweis darauf, dass häufig zitierte Entscheidungen anderer Bundesgerichte womöglich auch Entscheidungen des Bundesverfassungsgerichtes zitieren. Unter der Beachtung, dass das Bundesverfassungsgericht als hierarchisch bedeutsamstes Gericht gilt, wäre dies durchaus eine plausible Hypothese, welche in zukünftigen Untersuchungen validiert werden kann. In der Tabelle 6.3 sind zudem die Entscheidungen *GrS 4/82* und *GrS 2/92* des sog. großen Senates des Bundesfinanzhofes aufgeführt. Dieser wird grundsätzlich nur aufgerufen, um Fragen von grundsätzlicher Bedeutung zu klären, sowie die Rechtsprechung in Steuersachen zu vereinheitlichen. Das Verfahren ordnet somit auch Entscheidungen mit einer übergreifenden Bedeutung zentral ein.

Im Rahmen der netzwerktechnischen Untersuchungen auf der Mikroebene, können auch die Positionszentralitäten untersucht werden. Für das Netzwerk der Arbeit wurde für das Errechnen der Zwischenzentralität ein Algorithmus mit der Laufzeit  $O(nm + n^2 \log n)$  verwendet, wobei  $m$  die Anzahl der Kanten und  $n$  die Anzahl der Knoten darstellt.<sup>1</sup> Das Netzwerk enthielt zum Zeitpunkt der Auswertungen insgesamt 120.292 Knoten und 317.661 Kanten. Die Größe des Netzwerkes führte dazu, dass der Algo-

<sup>1</sup>vgl. [Bra01] S. 163

rithmus aufgrund der hohen Menge an Daten nach von fünf Stunden Berechnung nicht terminierte, weswegen für diese Berechnungen nun keine Messergebnisse vorliegen. Die Berechnung von kleineren Graphen, z.B. generierten Clustern terminierte korrekt, weswegen der Algorithmus grundsätzlich korrekte Ergebnisse liefert. Die Berechnung der Nähezentralität terminierte, allerdings wurde für jeden Knoten lediglich der Wert 0.0 berechnet. Es liegt hier die Vermutung nahe, dass der verwendete Algorithmus der Bibliothek JGraphT, welcher die Werte mit dem Datentyp *Double* berechnet, den errechneten Wert nicht erfassen kann, da dieser zu niedrig ist. Um die Vermutungen zu überprüfen, soll in einer zukünftigen Untersuchung überprüft werden, warum hier keine Werte errechnet werden konnten. Grundsätzlich zeigen Knoten mit einer hohen Zwischenzentralität gezeigt, dass diese z.B. getrennte Teile eines Netzwerkes verbinden. Anhand von Entscheidungen mit einer hohen Zwischenzentralität kann überprüft werden, inwieweit diese Entscheidung dazu beigetragen hat, dass verschiedene Gruppierungen im Netzwerk miteinander verbunden sind. Diese Entscheidungen stellen damit im Prinzip einen Vermittler dar. Eine hohe Nähezentralität einer Entscheidung zeigt wiederum, dass diese sehr nah mit vielen anderen Entscheidungen verknüpft ist.

Eine Untersuchung der Mesoebene, welche die regionalen Strukturen eines Netzwerkes untersucht, wurde hier in einem Blindtestverfahren durchgeführt. Dafür wurden Ausschnitte der von dem Algorithmus Chinese Whispers generierten Cluster von einem Experten untersucht. Im Folgenden sollen an zwei Beispielen die Ergebnisse dieses Verfahrens demonstriert werden. Die Tabelle 6.4 zeigt einen Cluster, welcher von einem

Aktenzeichen	Themengebiet
1 BvR 2074/05	Automatisierte Kennzeichenerfassung, Datenschutz, Persönlichkeitsrecht
3 AZR 96/18	Betriebliche Altersversorgung Anpassung Versorgungszusagen
3 AZR 290/18	Betriebliche Altersversorgung, Anpassung, Auslegung einer Versorgungszusage
3 AZR 483/86	Betriebsvereinbarung und Ruhegeld
3 AZR 557/03	Eingriff in die dem Ruheständler zugesagte Witwenrente
11 BV 28/09	Wirksamkeit einer Betriebsratswahl und eines Zuordnungstarifvertrags
3 AZR 105/18	Betriebliche Altersversorgung, Anpassung, Auslegung einer Versorgungszusagen
4 Sa 67/17	Betriebsrentenanpassung
3 AZR 483/86	Betriebsvereinbarung und Ruhegeld
11 BV 28/09	Wirksamkeit einer Betriebsratswahl

Tabelle 6.4: Clusterevaluation: Betriebsrenten

Experten qualitativ analysiert wurde. An diesem Beispiel ist zu sehen, dass sich fast alle Entscheidungen des Clusters thematisch durch die Thematik *Betriebliche Altersvorsorge* bzw. *Betriebsrente* überschneiden. Die Entscheidung des Bundesverfassungsgerichtes (*1 BvR 2074/05*) wirkt allerdings wie ein Fremdkörper im Cluster, da sie inhaltlich

## 6 Ergebnisse

keinen Zusammenhang zum Rest des Clusters aufweist. Hier könnte es sich möglicherweise um ein Artefakt des verwendeten Algorithmus' handeln. Ein weiteres Beispiel eines generierten Clusters ist in Tabelle 6.5 zu sehen. Es ist zu erkennen, dass die Entscheidungen inhaltlich zusammenpassen, da sich alle Entscheidungen mit der *Haftung für Kapitalanlagen* beschäftigen. Interessanterweise finden sich im Cluster zudem unterschiedliche Anlageformen, welche normalerweise nicht dem gleichen Rechtsgebiet unterfallen, aber inhaltlich zusammengehören. Die generierten Cluster des Netzwerkes

Aktenzeichen	Themengebiet
XI ZR 4/88	Aufklärungspflicht durch Anlagevermittler, Bauherrenmodell
II ZR 89/91	Kapitalanlegerhaftung bei Publikumskommanditgesellschaft
V ZR 136/03	Schadensersatzansprüche wegen Schmiergeldabrede
XI ZR 37/03	Widerruflichkeit von Verbraucherkrediten zur Finanzierung einer Eigentumswohnung als Kapitalanlage
XI ZR 112/05	Schadensersatz wegen vorvertraglicher Pflichtverletzung bei Erwerb und Finanzierung einer Eigentumswohnung
VI ZR 246/81	Schadensersatzpflicht wegen § 826 BGB beim Erwerb eines Grundstücks
II ZR 17/17	Schadensersatz aufgrund unrichtiger Prospektangaben
XI ZR 232/09	Bankenhaftung bei einer finanzierten Beteiligung an einem geschlossenen Immobilienfond
II ZR 193/11	Vorvertragliche Aufklärungspflicht des Treuhandkommanditisten
XI ZR 132/07	Bankenhaftung bei finanzierter Kapitalanlage

Tabelle 6.5: Clusterevaluation: Haftung für Kapitalanlagen

können außerdem innerhalb des Frontends dieser Arbeit interaktiv untersucht werden. Über den Menüpunkt *Cluster*, können dabei alle generierten Cluster in der Anwendung dargestellt werden. Werden für das Netzwerk neue Cluster generiert, aktualisiert sich die Website automatisch, sodass immer die aktuellsten Cluster angesehen werden können. Die Abbildung 6.1 zeigt exemplarisch die Ansicht eines generierten Clusters. Innerhalb der Cluster Ansicht sind die Knoten farblich markiert, um die jeweilige Gerichtsform zu kennzeichnen. Ist ein Knoten Grau, handelt es sich um eine Entscheidung, welche nicht Teil des Korpus ist. Um genauere Informationen zu einer Entscheidung zu erhalten, kann der Knoten angeklickt werden. Anschließend wird der Nutzer auf die Detailseite der Entscheidung weitergeleitet und kann alle vorhandenen Informationen dieser Entscheidung einsehen. Diese Ansicht ist allerdings nur für Entscheidungen möglich, welche Teil des Korpus sind.

Zuletzt sollen nun die Ergebnisse der Klassifikation des Revisionsausganges demonstriert werden. Das Modell wurde dafür auf einen Trainingsdatensatz von 3.724 Sätzen in insgesamt 20 Iterationen trainiert. Anschließend wurde eine Evaluation auf 1.242 Sätze

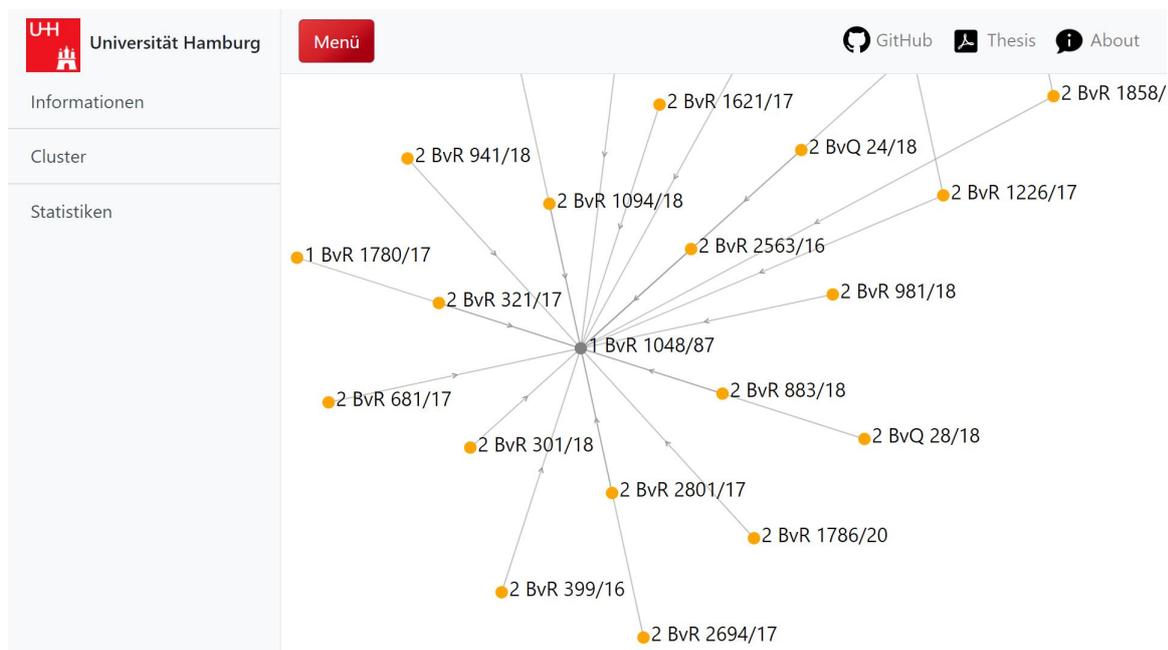


Abbildung 6.1: Netzwerkansicht eines generierten Clusters

ausgeführt. Die Ergebnisse des Verfahrens können in Tabelle 6.6 angesehen werden. Der Transformer kann hier mit einem F-Score von 0.96 sehr präzise Ergebnisse bestimmen. In der Tabelle sind zudem die Scores für die verschiedenen Klassen notiert. Da

Klasse	Precision	Recall	$F_1$	Anzahl
Positiv	0.97	0.96	0.96	338
Neutral	0.00	0.00	0.00	1
Negativ	0.96	0.96	0.96	305
Irrelevant	0.96	0.97	0.97	598
Gesamt	0.96	0.96	0.96	1.242

Tabelle 6.6: Evaluation der Klassifikation

der Trainingsdatensatz und der Testdatensatz zufällig bestimmt wurden, findet sich im Testdatensatz lediglich ein neutraler Satz. Hier kann in zukünftigen Untersuchungen ein Datensatz mit einer höheren Anzahl an neutralen Entscheidungen herangezogen werden, um die Messergebnisse zu verbessern.

Abschließend wird im Fazit eine Resümee über die durchgeführte Untersuchung vorgenommen.



## 7 Fazit

Diese Arbeit fungiert als Beispiel dafür, wie in dem Gebiet der Rechtswissenschaft, eine Untersuchung mit netzwerk- und sprachtechnologischen Techniken umgesetzt werden kann. Es wurde im Verlauf dieser Arbeit deutlich, welche Aufgaben der juristischen Netzwerkforschung zugrunde liegen und wie diese in ihren übergeordneten Forschungskontext eingeordnet werden kann. Dabei wurde sowohl auf die theoretischen Grundlagen der Netzwerkforschung Bezug genommen und es wurden verschiedene Verfahren vorgestellt, welche bei der automatisierten und systematischen Untersuchung von Netzwerken verwendet werden können. Zusätzlich dazu, wurde anhand von modernen Verfahren wie LSTM und Transformer Modellen gezeigt, an welchen Stellen Algorithmen der Sprachtechnologie und des maschinellen Lernens im Recht Anwendung finden. Anhand einer Softwareanwendung wurde skizziert, welche Aufgaben eine Infrastruktur einer solchen datengetriebenen Untersuchung erfüllen muss. Ferner wurde durch die entwickelte Webanwendung ein Ansatz vorgestellt, um juristische Netzwerke auch Adressatenkreisen näherzubringen, welche über ein Fachpublikum hinaus gehen.

Die Arbeit verfolgte das Ziel, anhand verschiedener Verfahren die Aussagekraft eines Entscheidungsnetzwerkes der obersten Gerichtsbarkeiten sowie des Bundesverfassungsgerichts und des Bundespatentgerichts zu erproben. Die Ergebnisse der Arbeit zeigen, dass die zugrunde liegenden Verfahren ein grundsätzlich valider Weg sind, die Möglichkeiten und Grenzen von juristische Netzwerkforschung zu erkennen. Es zeigt sich, dass in den hier untersuchten Bereichen mit einem Netzwerk aus Gerichtsentscheidungen durchaus Ergebnisse gewonnen werden können, in welchen die Rechtsprechung zu unterschiedlich großen Teilen durchaus wiederzuerkennen ist. Hierbei beeinflussen allerdings die Entscheidungen und Auslegungen verschiedener Begrifflichkeiten stark die Ergebnisse eines solchen Experiments, was bei der Deutung dieser berücksichtigt werden sollte. Die sprachtechnologische Untersuchung demonstriert dazu, dass die Rechtswissenschaft von präzisen modernen Methoden der Informatik profitieren kann. Mit dem transformerbasierten Modell konnte verdeutlicht werden, dass bei der Klassifikation des Revisionsausganges einer Gerichtsentscheidung mit State-of-the-Art Methoden auch mit einer geringen Menge an Trainingsdaten eine sehr hohe Präzision von  $F_1 = 0.96$  erreicht werden kann.

In einer weiteren Untersuchung, welche basierend auf den Ergebnissen dieser Arbeit stattfinden soll, können weitere Zusammenhänge betrachtet werden. So kann beispiels-

## 7 Fazit

weise die Abhängigkeit von Richtern zu der Textlänge von Entscheidungen genauer untersucht werden. Welche Richter verfassen besonders lange Urteile? Ebenso können Korrelationen von Richtern und anderen Metadaten der Entscheidungen untersucht werden. Bei welchen an Entscheidungen beteiligten Richtern tauchen häufig abweichende Meinungen auf? Gibt es Richter, welche auffällig häufig an zentralen Entscheidungen mitgewirkt haben? Eine weitere Möglichkeit ist die Betrachtung von Korrelationen zwischen dem Revisionserfolg und Richtern. Welche Richter verabschieden am häufigsten Entscheidungen, bei denen die Revision erfolgreich ist? Die weitere Forschung soll außerdem dazu genutzt werden, um auf die in dieser Untersuchung gewonnen Erkenntnisse aufzubauen und die erprobten Ansätze zu verbessern. Hierfür besteht exemplarisch die Möglichkeit, Daten für weitere Fundstellen zu sammeln, um mehr Zitate korrekt abbilden zu können. Es soll auch ein transformerbasiertes Modell erprobt werden, um Richter mit einer noch höheren Präzision aus einem Entscheidungstext extrahieren zu können. Ergebnisse, welche in dieser Arbeit nicht ermittelt werden konnten, sollen in einer zukünftigen Untersuchung aufgearbeitet werden

Es würde sich zudem auch anbieten weitere Metriken im Bereich der Netzwerkforschung zu evaluieren. So kann beispielsweise untersucht werden, ob es der Graph zusammenhängend ist, sich also die einzelnen Gerichtsbarkeiten untereinander zitieren, oder nicht. Auch eine Erweiterung der Datenbasis kann in Betracht gezogen werden. Hierfür können weitere Datenquellen evaluiert werden, um an mehr Entscheidungsdaten zu gelangen. Des Weiteren besteht die Möglichkeit die Analysen dieser Arbeit zukünftig durch ein Word2Vec-Modell zu erweitern. Hierfür wurde bereits zu den Autoren des Juristischen Referenzkorpus (JuReKo) Kontakt aufgenommen, welches weitere 370.000 deutsche Entscheidungen beinhaltet.

Der technische Fortschritt eröffnet der Rechtswissenschaft durch moderne Methoden neue Türen und erschafft Möglichkeiten, um systematisch an neue Informationen zu gelangen und bestehende Verfahren zu unterstützen oder sogar zu automatisieren. Dabei wird an vielen Stellen deutlich, dass die Fachrichtungen Informatik und Rechtswissenschaft mit der zunehmenden Digitalisierung immer weiter zusammenwachsen. Dass die Zusammenarbeit beider Disziplinen in Zukunft eine noch höhere Bedeutung erfährt, ist sehr wahrscheinlich. Allerdings wird auch die Entwicklung von rechtswissenschaftlichen Datenquellen interessant sein, um Untersuchungen in der quantitativen Rechtswissenschaft zukünftig noch präziser und umfangreicher durchführen zu können.

# Literaturverzeichnis

- [Alp20] ALPAYDIN, ETHEM: *Introduction to machine learning*. MIT press, 2020.
- [ATPPL16] ALETRAS, NIKOLAOS, DIMITRIOS TSARAPATSANIS, DANIEL PREOTIUC-PIETRO VASILEIOS LAMPOS: *Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective*. PeerJ Comput. Sci., 2:e93, 2016. <https://doi.org/10.7717/peerj-cs.93>.
- [Bec20] BECKEDORF, JANIS ET AL.: *Analyzing high volumes of German court decisions in an interdisciplinary class of law and computer science students - LSIA*. COMPUTATIONAL LEGAL STUDIES: THE PROMISE AND CHALLENGE OF DATA-DRIVEN LEGAL RESEARCH, 2020. (forthcoming).
- [Bie06] BIEMANN, CHRIS: *Chinese Whispers: An Efficient Graph Clustering Algorithm and Its Application to Natural Language Processing Problems*. *Proceedings of the First Workshop on Graph Based Methods for Natural Language Processing*, TextGraphs-1, 73–80, USA, 2006. Association for Computational Linguistics.
- [BMPB15] BENIKOVA, DARINA, SEID MUHIE, YIMAM PRABHAKARAN SANTHANAM CHRIS BIEMANN: *C.: GermaNER: Free Open German Named Entity Recognition Tool*. In: *Proc. GSCL-2015*, 2015.
- [Bra01] BRANDES, ULRİK: *A faster algorithm for betweenness centrality*. *Journal of mathematical sociology*, 25(2):163–177, 2001.
- [Bra10] BRANDES, ULRİK: *Graphentheorie*, 345–353. VS Verlag für Sozialwissenschaften, Wiesbaden, 2010.
- [CEE+09] CARSTENSEN, KAI-UWE, CHRISTIAN EBERT, CORNELIA EBERT, SUSANNE JEKAT, HAGEN LANGER RALF KLABUNDE: *Computerlinguistik und Sprachtechnologie: Eine Einführung*. Springer-Verlag, 2009.

Literaturverzeichnis

- [CLLM20] CLARK, KEVIN, MINH-THANG LUONG, QUOC V LE CHRISTOPHER D MANNING: *Electra: Pre-training text encoders as discriminators rather than generators*. arXiv preprint arXiv:2003.10555, 2020.
- [CN16] CHIU, JASON P.C. ERIC NICHOLS: *Named Entity Recognition with Bidirectional LSTM-CNNs*. Transactions of the Association for Computational Linguistics, 4:357–370, 2016.
- [Cou18] COUPETTE, CORINNA UND FLECKNER, ANDREAS M.: *Quantitative Rechtswissenschaft (Quantitative Legal Studies)*. JuristenZeitung (JZ), 73(8):379–389, 2018. <https://ssrn.com/abstract=3377384>.
- [Cou19] COUPETTE, CORINNA: *Juristische Netzwerkforschung - Modellierung, Quantifizierung und Visualisierung relationaler Daten im Recht*. Mohr Siebeck, Tübingen, 2019.
- [Dal19] DALE, ROBERT: *Law and word order: NLP in legal tech*. Natural Language Engineering, 25:211–217, 01 2019.
- [DJ10] DAVID, EASLEY KLEINBERG JON: *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cambridge University Press, USA, 2010.
- [DL14] DERLÉN, MATTIAS JOHAN LINDHOLM: *Goodbye van Gend en Loos, Hello Bosman? Using Network Analysis to Measure the Importance of Individual CJEU Judgments*. European Law Journal, 20(5):667–687, 2014.
- [EdCMMY<sup>+</sup>16] CASTILHO, RICHARD ECKART DE, ÉVA MÚJDRICZA-MAYDT, SEID MUHIE YIMAM, SILVANA HARTMANN, IRYNA GUREVYCH, ANETTE FRANK CHRIS BIEMANN: *A Web-based Tool for the Integrated Annotation of Semantic and Syntactic Structures*. Proceedings of the Workshop on Language Technology Resources and Tools for Digital Humanities (LT4DH), 76–84, Osaka, Japan, 2016. The COLING 2016 Organizing Committee.
- [EGS<sup>+</sup>17] ENGST, BENJAMIN, THOMAS GSCHWEND, NILS SCHAKS, SEBASTIAN STERNBERG CAROLINE WITTIG: *Zum Einfluss der Parteinähe auf das Abstimmungsverhalten der Bundesverfassungsrichter – eine quantitative Untersuchung*. JuristenZeitung, 72:816–826, 09 2017.
- [Fir12] FIRDHOUS, MOHAMED: *Automating Legal Research through Data Mining*, 2012.

- [FJS<sup>+</sup>07] FOWLER, JAMES H., TIMOTHY R. JOHNSON, JAMES F. SPRIGGS, SANGICK JEON PAUL J. WAHLBECK: *Network Analysis and the Law: Measuring the Legal Importance of Precedents at the U.S. Supreme Court*. *Political Analysis*, 15(3):324–346, 2007.
- [GBB11] GLOT, XAVIER, ANTOINE BORDES YOSHUA BENGIO: *Deep Sparse Rectifier Neural Networks*. GORDON, GEOFFREY, DAVID DUNSON MIROSLAV DUDÍK (): *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, 15 *Proceedings of Machine Learning Research*, 315–323, Fort Lauderdale, FL, USA, 11–13 Apr 2011. JMLR Workshop and Conference Proceedings.
- [GMH13] GRAVES, A., A. MOHAMED G. HINTON: *Speech recognition with deep recurrent neural networks*. *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, 6645–6649, 2013.
- [GP17] GAO, BOLIN LACRA PAVEL: *On the properties of the softmax function with application in game theory and reinforcement learning*. arXiv preprint arXiv:1704.00805, 2017.
- [GS12] GELTER, MARTIN MATHIAS SIEMS: *Networks, Dialogue or One-Way Traffic? An Empirical Analysis of Cross-Citations between Ten of Europe's Highest Courts*. 8(2):88–99, 2012. <https://ssrn.com/abstract=2063598>.
- [Har20] HARTUNG, DIRK: *Quantitative Legal Research in Germany*. (forthcoming), 2020.
- [HS97] HOCHREITER, S. J. SCHMIDHUBER: *Long Short-Term Memory*. *Neural Computation*, 9:1735–1780, 1997.
- [IL14] IWANOWSKI, SEBASTIAN RAINER LANG: *Diskrete Mathematik mit Grundlagen - Lehrbuch für Studierende von MINT-Fächern*. Springer-Verlag, Berlin Heidelberg New York, 2014.
- [JM08] JURAFSKY, DANIEL JAMES MARTIN: *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*, 2. 02 2008.
- [KB17] KINGMA, DIEDERIK P. JIMMY BA: *Adam: A Method for Stochastic Optimization*, 2017.

Literaturverzeichnis

- [KCBH20] KATZ, DANIEL MARTIN, CORINNA COUPETTE, JANIS BECKENDORF DIRK HARTUNG: *Complex societies and the growth of the law*. Scientific Reports, 10, 2020.
- [KS08] KATZ, DANIEL MARTIN DEREK STAFFORD: *Hustle and Flow: A Social Network Analysis of the American Federal Judiciary*. Political Methods: Computational eJournal, 2008.
- [LFM<sup>+</sup>18] LETTIERI, NICOLA, SEBASTIANO FARO, DELFINA MALANDRINO, ARMANDO FAGGIANO MARGHERITA VESTOSO: *Network, Visualization, Analytics. A Tool Allowing Legal Scholars to Experimentally Investigate EU Case Law: AICOL International Workshops 2015-2017: AICOL-VI@JURIX 2015, AICOL-VII@EKAW 2016, AICOL-VIII@JURIX 2016, AICOL-IX@ICAIL 2017, and AICOL-X@JURIX 2017, Revised Selected Papers*, 10791, 543–555. 01 2018.
- [McC80] MCCARTY, L. THORNE: *The TAXMAN Project: Towards a Cognitive Theory of Legal Argument*, 23–43. 01 1980.
- [MR11] MIHALCEA, RADA DRAGOMIR RADEV: *Graph-based Natural Language Processing and Information Retrieval*. Cambridge University Press, 2011.
- [Nay18] NAY, JOHN: *Natural Language Processing and Machine Learning for Law and Policy Texts*. 2018. <https://ssrn.com/abstract=3438276>.
- [NS00] NOTO, M. H. SATO: *A method for the shortest path search by extended Dijkstra algorithm. Smc 2000 conference proceedings. 2000 ieee international conference on systems, man and cybernetics. 'cybernetics evolving to systems, humans, organizations, and their complex interactions' (cat. no.0, 3, 2316–2320 vol.3, 2000*.
- [PBMW99] PAGE, LAWRENCE, SERGEY BRIN, RAJEEV MOTWANI TERRY WINOGRAD: *The PageRank citation ranking: Bringing order to the web.* , Stanford InfoLab, 1999.
- [RHS<sup>+</sup>18] RUPPERT, EUGEN, DIRK HARTUNG, PHILLIP SITTIG, TJORBEN GSCHWANDER, LENNART RÖNNEBURG, TOBIAS KILLING CHRIS BIEMANN: *LawStats - Large-Scale German Court Decision Evaluation Using Web Service Classifiers*. CD-MAKE, 2018.
- [RR09] RATINOV, LEV-ARIE D. ROTH: *Design Challenges and Misconceptions in Named Entity Recognition*. CoNLL, 2009.

- [Sus86] SUSSKIND, RICHARD E.: *EXPERT SYSTEMS IN LAW: A JURISPRUDENTIAL APPROACH TO ARTIFICIAL INTELLIGENCE AND LEGAL REASONING*. The Modern Law Review, 49(2):168–194, 1986.
- [TYS+20] TANG, PIN, PINLI YANG, YUANG SHI, YI ZHOU, FENG LIN YAN WANG: *Recognizing Chinese Judicial Named Entity using BiLSTM-CRF*, 2020.
- [VSP+17] VASWANI, ASHISH, NOAM SHAZEER, NIKI PARMAR, JAKOB USZKOREIT, LLION JONES, AIDAN N. GOMEZ, L. KAISER ILLIA POLOSUKHIN: *Attention is All you Need*. NIPS, 2017.
- [Wal18] WATTL, BERNHARD ERNST: *Semantic Analysis and Computational Modeling of Legal Documents*. , Technische Universität München, 2018.
- [WBS+17] WATTL, BERNHARD, GEORG BONCZEK, ELENA SCEPANKOVA, JOERG LANDTHALER FLORIAN MATTHES: *Predicting the Outcome of Appeal Decisions in Germany's Tax Law*. 89–99, 07 2017.
- [WRK11] WINKELS, RADBOUD, JELLE RUYTER KROESE: *Determining Authority of Dutch Case Law*. Frontiers in Artificial Intelligence and Applications, 235, 12 2011.
- [YB19] YADAV, VIKAS STEVEN BETHARD: *A Survey on Recent Advances in Named Entity Recognition from Deep Learning models*, 2019.



### Eidesstattliche Erklärung

Hiermit versichere ich an Eides statt, dass ich die vorliegende Arbeit im Bachelorstudiengang Software-System-Entwicklung selbstständig verfasst und keine anderen als die angegebenen Hilfsmittel — insbesondere keine im Quellenverzeichnis nicht benannten Internet-Quellen — benutzt habe. Alle Stellen, die wörtlich oder sinngemäß aus Veröffentlichungen entnommen wurden, sind als solche kenntlich gemacht. Ich versichere weiterhin, dass ich die Arbeit vorher nicht in einem anderen Prüfungsverfahren eingereicht habe und die eingereichte schriftliche Fassung der auf dem elektronischen Speichermedium entspricht.

Hamburg, den 23.12.2020

  
Lennart Rönneburg

### Veröffentlichung

Ich stimme der Einstellung der Arbeit in die Bibliothek des Fachbereichs Informatik zu.

Hamburg, den 23.12.2020

  
Lennart Rönneburg